UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/989,920	11/21/2001	Roberto A. Macina	DEX-0291	2846

7590 12/09/2004		EXAMINER
Licata & Tyrrell P.C.		YU, MISOOK
66 East Main Street		
Marlton, NJ 08053	ART UNIT	PAPER NUMBER
	1642	

DATE MAILED: 12/09/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

		Application No.	Applicant(s)
Office Action Summary		09/989,920	MACINA ET AL.
		Examiner	**Art Unit**	
		MISOOK YU, Ph.D.	1642	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
 Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on *27 September 2004*.

2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *1-17* is/are pending in the application.

   4a) Of the above claim(s) *6,10-14 and 16* is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) *1-5,7-9,15 and 17* is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.

10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.

   Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

   Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   a) ☐ All b) ☐ Some * c) ☐ None of:

   1. ☐ Certified copies of the priority documents have been received.

   2. ☐ Certified copies of the priority documents have been received in Application No. _____.

   3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

   * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
   Paper No(s)/Mail Date *10/29/04,09/27/02*.

4) ☐ Interview Summary (PTO-413)
   Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☒ Other: *Exhibits A-D*.

DETAILED ACTION

Election/Restrictions

Applicant's election with traverse of group 100 encompassing claims 1-5, 7-9, 15,

and 17 partially, relating SEQ ID NO:100 nucleic acid molecule, and the nucleic acid

encoding the protein encoded by the nucleic acid molecules in the reply filed on

09/27/2004 is acknowledged. The traversal is on the ground(s) that searching of the

nucleic acids, polypeptide, and antibody all together would not put a serious burden on

the Examiner, and the restriction groups are not independent or distinct. Applicant also

argues that MPEP 803.04 says that up to 10 nucleotide sequences are examined in a

single application. These arguments have been fully considered but found

unpersuasive for the following reasons. First, all the nucleic acid sequences in the

instant application are a large sequences, for example, the elected SEQ ID NO:100 is

close to 3 kb, which requires search in sequence databases and non-patent literature.

Each of the different nucleotides sequences appear to be either cDNA, or genomic DNA

from the different human chromosomes without any relationship to each other.

Examination of up to 10 nucleotide sequences applies when the search burden of up to

10 nucleotide sequences is not serious on the examiner or the office's available

resources, for example the use of computer time. In the instant cases, each SEQ ID

NO is a large sequences and search of one application up to 10 nucleotides sequences

would put a serious burden on the Office.

As for the examination of nucleic acids along with the protein encoded by the

nucleic acid and the antibody binding to the protein, the polypeptide groups, and

antibody groups are different inventions. Polypeptides, which are composed of amino

acids, and polynucleotides, which are composed of purine and pyrimidine units, are

structurally distinct molecules; any relationship between a polynucleotide and

polypeptide is dependent upon the information provided by the nucleic acid sequence

open reading frame as it corresponds to the primary amino acid sequence of the

encoded polypeptide. In the present claims, a polynucleotide of group I does not

necessarily encode a polypeptide. In fact, the elected SEQ ID NO:100 does not encode

any protein in the instant application. Alignment of instant SEQ ID NO:100 against the

protein sequences i.e. SEQ ID NOs 165-284 reveals none of the protein is encoded by

instant SEQ ID NO:100. Note Exhibit A (total 20 pages). Further, the nucleic acid

molecule of claim 1(c) is complementary to the coding sequence. If SEQ ID NO:100 is a

coding sequence, and then the complementary sequence would not encode any

polypeptide because the complimentary sequence of a human cDNA does not usually

encode any protein. The instant specification does not teach any evidence to the

contrary. There is search burden also in the non-patent literature. Prior to the

concomitant isolation and expression of the sequence of interest there may be journal

articles devoted solely to polypeptides which would not have described the

polynucleotide. Similarly, there may have been "classical" genetics papers, which had

no knowledge of the polypeptide but spoke to the gene. Searching, therefore is not

coextensive. The polypeptide of groups and the antibody groups are patentably distinct

for the following reasons: While the inventions of both groups are polypeptides, in this

instance the polypeptide is a single chain molecule, whereas antibody encompasses

IgG which comprises 2 heavy and 2 light chains containing constant and variable regions, and including framework regions which act as a scaffold for the 6 complementarity determining regions (CDRs) that function to bind an epitope. Therefore they are structurally distinct molecules; any relationship between a polypeptide and an antibody is dependent upon the correlation between the scope of the polypeptides that the antibody binds and the scope of the antibodies that would be generated upon immunization with the polypeptide.

The polynucleotide and the antibody group are patentably distinct. An antibody for example, IgG molecule which comprise 2 heavy and 2 light chains containing constant and variable regions, and including framework regions which act as a scaffold for the 6 complementarity determining regions (CDRs). Polypeptides, such as the antibody of group II which are composed of amino acids, and polynucleotides, which are composed of nucleic acids, are structurally distinct molecules; any relationship between a polynucleotide and polypeptide is dependent upon the information provided by the nucleic acid sequence open reading frame as it corresponds to the primary amino acid sequence of the encoded polypeptide. In the present claims, a polynucleotide in the nucleic acid groups will not encode an antibody in antibody groups, and the antibody of group III cannot be encoded by a polynucleotide of group I. Therefore the antibody and polynucleotide are patentably distinct.

The requirement is still deemed proper and is therefore made FINAL.

Claims 6, and 10-14, and 16 are withdrawn from further consideration pursuant to 37 CFR 1.142(b), as being drawn to a nonelected invention, there being no allowable generic or linking claim.

Claims 1-17 are pending, and claims 1-5, 7-9, 15 and 17 are examined as to the extent they are drawn to the elected invention of SEQ ID NO:100 or related nucleic acid sequence.

Specification

The disclosure is objected to because it contains an embedded hyperlink and/or other form of browser-executable code, for example at pages 54, 61. Applicant is kindly requested to go through the entire specification very carefully to see if there is any other embedded hyperlink and/or other form of browser-executable code. Applicant is required to delete all embedded hyperlink and/or other form of browser-executable code(s) in the specification. See MPEP § 608.01.

Claim Objections

Claims 1-5, 7-9, 15, and 17 are objected to because of the following informalities: the claims are still drawn to multiple inventions. The claims have not been amended to reflect the election. Appropriate correction is required.

Claim Rejections - 35 USC § 101

Claims 1-5, 7-9, 15, and 17 are rejected under 35 U.S.C. 101 because the claimed invention is not supported by either a substantial utility, or a well established utility.

Claim 1-5, 7-9, 15, and 17 are interpreted as drawn to SEQ ID NO:100 nucleic

acid, kit containing said nucleic acid, vector containing said nucleic acid, host cell

containing said vector, method of producing protein using said host cells. The

specification speculates that SEQ ID NO:100 might have utilities in making protein,

making antibody, diagnostic and staging assays for lung cancer or non-cancerous

diseases (at pages 93-103) or detecting a risk of cancer or presence of cancer (claim

15), method of identifying lung tissue (page 103), method of producing and modifying

lung tissue such as making an artificial lung (at pages 104-105), pharmaceutical (page

105) in gene therapy and antisense therapy (pages 111-113).

These utilities are not considered to substantial enough because neither the

specification nor any art of record teaches what the biological activities of SEQ ID

NO:100 are. The specification at page 6, lines 12-27 teaches that the disclosed nucleic

acids are lung specific. An assay to tell whether one has lung i.e. lung tissue typing is

not considered a substantial enough utility. The specification asserts that the differential

expression of the sequence is used for lung cancer detection. However, the

specification does not teach whether the claimed nucleic acid is under-expressed or

overexpressed in lung cancer. The specification does not teach a relationship to any

specific disease or establish any involvement SEQ ID NO:100. The specification does

not teach which protein is encoded by SEQ ID NO:100, let alone substantial or specific

use for it. None of the protein sequences disclosed in the instant application is encoded

by instant SEQ ID NO:100. Note the Exhibit A. Making and purifying the protein

encoded by SEQ ID NO:100 does not lead to a substantial use of the claimed invention

because neither the specification nor the art appears to know what the structure of the

protein encoded by the claimed invention. In fact, Genbank Accession No. AC079988

(gi:18873965, #DI of IDS filed on 10/29/2004) teach a genomic DNA i.e. human BAC

clone RP11-795C1 (form chromosome 2) having 98.7 % sequence identity to instant

SEQ ID NO:100. Note the sequence alignment of instant SEQ ID NO:100 against

Genbank Accession No. AC079988 (Exhibit B). Further, Genbank Accession No.

AC079988 teach that this clone is from RPCI-11 human BAC library prepared from the

blood of one male donor as disclosed in The Sanger Center and The Washington

University Genome Sequencing Center (DC of IDS filed on 10/29/2004, 1998, Genome

Research, vol. 8, pages 1097-1108. The Sanger Center and The Washington

University Genome Sequencing Center teaches that the library was constructed for

sequencing human genome, and it is not cDNA library. Therefore, SEQ ID NO:100 is

most likely a genomic sequence, not specific for lung only but present in every cell of

human body that contains chromosome 2.

The asserted utilities as hybridization probes, antisense, the various assays

numerated in the instant application do not lead to substantial and credible uses of the

claimed invention due to unknown functions of the protein encoded by the claimed

invention. Nothing is specific to the sequences of the claimed invention for all of the

various probe uses. Any nucleic acid can be used to, identify polymorphisms, map

chromosomes, type a tissue, make transgenic animals or knockout animals. The

specification does not have any substantial use for pharmaceutical compositions,

diagnostic assay, and methods of treatment because the specification does not teach

what disease(s) is caused by malfunction of the claimed invention or the protein

encoded by it.   Since EQ ID NO:100 does not have a substantial utility, or a well

established utility, a compound that binds to SEQ ID NO:100 does not have a

substantial utility, or a well established utility.

In Brenner v. Manson, 148 U.S.P.Q. 689 (Sup. Ct., 1966), a process of producing

a novel compound that was structurally analogous to other compounds which were

known to possess anti-cancer activity was alleged to be useful because the compound

produced thereby was potentially useful as an anti-tumor agent in the absence of

evidence supporting this utility. The court expressed the opinion that all chemical

compounds are "useful" to the chemical arts when this term is given its broadest

interpretation. However, the court held that this broad interpretation was not the

intended definition of "useful" as it appears in 35 U.S.C. § 101, which requires that an

invention must have either an immediately obvious or fully disclosed "real world" utility.

The instant claims are drawn to SEQ ID NO:100 which has undetermined function or

biological significance. Until some actual and specific activity can be attributed to the

nucleic acid or the protein encoded by the claimed invention encoding it, the claimed

invention is incomplete.

Claims 1-5, 7-9, and 17 are also rejected under 35 U.S.C. 112, first paragraph.

Specifically, since the claimed invention is not supported by either a substantial

asserted utility or a well established utility for the reasons set forth above, one skilled in

the art clearly would not know how to use the claimed invention.

Claim Rejections - 35 USC § 112

The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 1-5, 7-9, 15, and 17 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

The claims are interpreted as drawn to a genus of nucleic acid molecules with various of degrees of variations from SEQ ID NO:100, a genus of vectors containing said nucleic acid molecules, genus of host cells containing said vectors, a method of producing a genus of polypeptides using said host cells.

The applicable standard for the written description requirement can be found: MPEP 2163; University of California v. Eli Lilly, 43 USPQ2d 1398 at 1407; PTO Written Description Guidelines; Enzo Biochem Inc. v. Gen-Prove Inc., 63 USPQ2d 1609; Vas-Cath Inc. v. Mahurkar, 19USPQ2d 1111; and University of Rochester v. G.D. Searle & Co., 69 USPQ2d 1886 (CA FC 2004).

To provide adequate written description and evidence of possession of a claimed genus, the specification must provide sufficient distinguishing identifying characteristics of the genus. The factors to be considered include disclosure of complete or partial structure, physical and/or chemical properties, functional characteristics,

structure/function correlation, methods of making the claimed product, or any combination thereof. In this case, the only factor present in the claims is a partial structure in the form of percent identity or hybridization. There is not even identification of any associated function with the claimed genus of partial structures. Accordingly, in the absence of sufficient recitation of distinguishing identifying characteristics, the specification does not provide adequate written description of the claimed genus.

Vas-Cath Inc. v. Mahurkar, 19USPQ2d 1111, clearly states "applicant must convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession of the invention. The invention is, for purposes of the 'written description' inquiry, *whatever is now claimed.*" (See page 1117.) The specification does not "clearly allow persons of ordinary skill in the art to recognize that [he or she] invented what is claimed." (See Vas-Cath at page 1116). As discussed above, the skilled artisan cannot envision the detailed chemical structure of the encompassed genus of nucleic acid molecules, given that the specification has only described SEQ ID NO: 100. Therefore, only isolated nucleic acid comprising SEQ ID NO:100, but not the full breadth of the claim meets the written description provision of 35 U.S.C. §112, first paragraph.

Further, claim 9 is drawn to method of producing a polypeptide using the claimed nucleic acid. However, the specification does not teach the structure of the polypeptide the instant SEQ ID NO:100 encodes, thus failing to provide an adequate written description.

Claims 1-5, 7-9, 15, and 17 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the **enablement requirement**. The claim(s) contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

The factors considered when determining if the disclosure satisfies the enablement requirement and whether any necessary experimentation is "undue" include, but are not limited to: 1) nature of the invention, 2) state of the prior art, 3) relative skill of those in the art, 4) level of predictability in the art, 5) existence of working examples, 6) breadth of claims, 7) amount of direction or guidance by the inventor, and 8) quantity of experimentation needed to make or use the invention. *In re Wands*, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988).

This enablement rejection is made based on the interpretation of the claims as drawn to an isolated nucleic acid molecule comprising SEQ ID NO:100, a nucleic acid that selectively hybridizes, or at least 60 % sequence identity to instant SEQ ID NO:100 (for use in lung cancer detection (note pages 93-103, abstract) or use as a vaccine (note claim 17). Further, claim 8 is drawn to a host cell comprising the vector of claim 8. therefore claim as currently construed is broadly interpreted to encompass host cells, which are not isolated and are comprised within an organism. Thus, the claim encompass host cells that have been transfected with the vector of claims 8 that are comprised within a transgenic animal, including nonhuman or human animals and animals treated using gene therapy (note page 111 of the specification).

The specification does not teach whether SEQ ID NO:100 is over-expressed or

under expressed in lung cancer or any other lung disease, let alone a nucleic acid that

selectively hybridizes, or at least 60 % sequence identity to instant SEQ ID NO:100

being over-expressed or under-expressed in any lung disease including lung cancer.

The specification provides neither guidance on nor exemplification of how to correlate

the data presented in the specification with the ability to use SEQ ID NO:100 for the

assessment of cancer risk. In other words, the specification does not present any in

vivo data to correlate either detection of the nucleic acid or absence of the nuceic acid

to growth of any tumor. Tockman et al (Cancer Res., 1992, 52:2711s-2718s) teach

considerations necessary in bringing a lung cancer biomarker to successful clinical

application. Tockman et al teach that prior to the successful application of newly

described lung cancer markers, research must validate the markers against

acknowledged disease end points, establish quantitative criteria for marker

presence/absence and confirm marker predictive value in prospective population trials

(see abstract). Early stage markers of tumorigenicity have clear biological plausibility as

markers of preclinical cancer and if validated can be used for population screening (p.

2713s, col 1). The reference further teaches that once selected, the sensitivity and

specificity of the biomarker must be validated to a known (histology/cytology-confirmed)

cancer outcome. The essential element of the validation of an early detection marker is

the ability to test the marker on clinical material obtained from subjects monitored in

advance of clinical cancer and link those marker results with subsequent histological

confirmation of disease. This irrefutable link between antecedent marker and

subsequent acknowledged disease is the essence of a valid intermediate end point

marker (p. 2714, see Biomarker Validation against Acknowledged Disease End Points).

Clearly, prior to the successful application of newly described markers, markers must be

validated against acknowledged disease end points and the marker predictive value

must be confirmed in prospective population trials (p. 2716s, col 2). The specification

provides insufficient guidance, and provides no working examples of correlating in lung

cancer to either detection of SEQ ID NO:100 or to absence of SEQ ID NO:100, which

would provide guidance to one skilled in the art to use the claimed invention without

undue experimentation. Considering lack of examples and the limited teachings of the

specification, and unpredictability in the art, it is concluded that undue experimentation

would be required to practice the claimed invention. It is noted that law requires that the

disclosure of an application shall inform those skilled in the art how to make the alleged

discovery, not how to screen it for themselves.

As for claim 17, the claim is interpreted as drawn to DNA vaccine comprising

SEQ ID NO:100 or other related nucleic acids as described in the base claim. Roos et

al., (06/25/2004, Prostate, vol. 9999, pages 1-7) teach that DNA vaccine is used to

generate immunity against a specific antigen in this case tumor antigen, and the

consideration necessary for DNA vaccine to work is to put the DNA encoding the

antigen of interest into a vector such that the antigen is expressed in vivo (note page 2,

left column). The instant application does not even teach the structure of protein being

encoded by the claimed nucleic acid. Note the attached Exhibit A, showing none of the

numerous protein sequences in the instant application is being encoded by instant SEQ

ID NO:100. Since DNA vaccine requires testing of its efficacy in vivo (note page 3, right

column of Roos et al.), and the instant specification does not even teach whether instant

SEQ ID NO:100 encodes any protein, let alone any antigen that an animal can produce

immune response against it, it requires undue experimentation to use the instant

claimed vaccine. The specification provides insufficient guidance, and provides no

working examples of how to use SEQ ID NO:100, and its related nucleic acid

sequences as vaccine, which would provide guidance to one skilled in the art to use the

claimed invention without undue experimentation. Considering lack of examples and

the limited teachings of the specification, and unpredictability in the art, and broad

breath, it is concluded that undue experimentation would be required to practice the

claimed invention.

As for claim 8, the specification does not provide a sufficient amount of guidance,

direction, or exemplification to enable the skilled artisan to make or use host cells that

are comprised within a non-human transgenic animal. In addition, the specification

does not teach provide a sufficient amount of guidance, direction, and exemplification to

enable the skilled artisan to have a reasonable expectation of successfully producing

host cells within a living organism, which comprise the vectors of claims 8, by gene

transfer, or gene therapy. The art of gene therapy, i.e., the in vivo delivery genetic

information to targeted cells within a body using naked DNA or viral vectors or by

reintroducing ex vivo modified host cells into the body, is still in its infancy. Moreover,

the art is highly unpredictable and its successful application has been hindered by

numerous limitations, which the specification does not remedy and would preclude the

skilled artisan from having a reasonable expectation of successfully making and using

the claimed invention without need of performing an undue amount of experimentation.

The specification does not teach method of gene therapy using the instantly claimed

nucleic acid. The art recognizes that gene therapy is not a trivial matter. Friedmann

(Scientific American, June 1997, pages 96-101), Verma and Somia (1997, Nature, vol.

389, pages 239-242), and Rubanyi (2001, Molecular Aspects of Medicine 22, pages

113-142) all teach that gene therapy art still faces major hurdle to overcome. For

example, the teachings of the specification have not overcome the problems with *in vivo*

delivery and expression. Verma et al. (*Nature* 1997, **389**: 239-242) teach that the

Achilles heel of gene therapy is gene delivery. Verma et al. state that the ongoing

problem is the inability to deliver genes efficiently and to obtain sustained expression.

Rubanyi at the abstract teaches that the prerequisite of successful gene therapy

includes "therapeutically suitable genes with a proven role in pathophysiology of the

disease". The instant specification fails at this first prerequisite because the

specification does not teach any therapeutically suitable genes with a proven role in

pathophysiology of the disease. Friedman summaries the current state of gene therapy

as "treating disease by providing needed gene remains a compelling idea, but clinical

and basic researchers still have much to do before gene therapy can live up to its

promise" (note the italicized headline at the top of page 96). The instant specification

does not teach a single technical problem being solved for gene therapy art. In view of

the preponderance of evidence establishing the state of the art, now and at the time the

application was filed, and the level of unpredictability associated therewith, in the

absence of a disclosure of an amount of guidance, direction, and exemplification that is

reasonably commensurate in scope with the claims, it appears that skilled artisan could

not make and use the claimed invention with a reasonable expectation of success

without having the need to perform an undue amount of experimentation. Amending

claim 8 to recite "isolated" before "host cell" can obviate this ground of rejection.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that

form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

(e) the invention was described in a patent granted on an application for patent by another filed in the United States before the invention thereof by the applicant for patent, or on an international application by another who has fulfilled the requirements of paragraphs (1), (2), and (4) of section 371(c) of this title before the invention thereof by the applicant for patent.

Claims 1-5, 7, 8, 15, and 17 are rejected under 35 U.S.C. 102(e) as being

anticipated by US Pat. 6,368,794 B1 (issued April 09, 2002, filed Jan. 15, 1999).

The claims are interpreted as drawn to an isolated nucleic acid molecule

comprising a nucleic acid that selectively hybridizes or at least 60 % sequence identity

to SEQ ID NO:100 (claim 1), wherein said nucleic acid molecule is a cDNA (claim 2),

genomic DNA (claim 3), a mammalian nucleic acid molecule (claim 4), a human nucleic

acid molecule (claim 5), in a vector (claim 7), in a host cell comprising said vector (claim

8), a kit comprising a means of for determining the presence of said nuclei acid, or said

nucleic acid molecule in a vaccine (claim 17).

US Pat. 6,368,794 B1 teaches SEQ ID NO:3 having 63.5 % sequence identity to instant SEQ ID NO:100. Note the sequence alignment of instant SEQ ID NO:100 against SEQ ID NO:3 of US Pat. 6,368,794 B1 (Exhibit C). The patent teaches that SEQ ID NO:3 is a cDNA obtained by reverse transcribing polyA mRNA (note columns 17-20, claim 1) from a human tissue, and also teaches method of detecting cancerous condition using a kit comprising a means to detect the nucleic acid of the base claim i.e., claim 1 of the patent (note claims 5-8).

Since the specification at paragraph [0034] defines "exon" refers to a nucleic acid sequence found in genomic DNA that is bioinformatically predicted and/or experimentally confirmed to contribute a contiguous sequence to a mature mRNA transcript, SEQ ID NO:3 of US Pat. 6,368,794 B1 is a genomic DNA. Note the SEQ ID NO:3 meets the definition of genomic DNA as defined in the instant application since it is transcribe from exons. Note SEQ ID NO:3 of the a patent is a mature mRNA.

As for claims 7, and 8, the patent (at column 15 under the heading "Isolation and sequencing of cDNA Clones") teaches that the clones disclosed in the patent are in plasmid in a host cell.

As for claim As for claim 17, the preamble recitation of "vaccine" is merely suggestive of an intended use and is not given patentable weight for purposes of comparing the claim with the prior art. The claim reads on nucleic acids *per se*.

Thus, US Pat. 6,368,794 B1 anticipates claims 1-5, 7, 8, 15, and 17.

Claims 1, 3-5, 7, 8, 15, and 17 are rejected under 35 U.S.C. 102(**a**) as being

anticipated by Genbank Accession No. AC079988 (gi:10800346 with public availability

date of 10/14/2000 online, IDS #DI filed on 10/29/04).

The claims are interpreted as drawn to an isolated nucleic acid molecule

comprising a nucleic acid that selectively hybridizes or at least 60 % sequence identity

to instant SEQ ID NO:100 (claim 1), wherein said nucleic acid molecule is genomic DNA

(claim 3), a mammalian nucleic acid molecule (claim 4), a human nucleic acid molecule

(claim 6), in a vector (claim 7), in a host cell comprising said vector (claim 8), kit

comprising a means to detect the nucleic acid of claim 1, or said nucleic acid molecule

in a vaccine (claim 17).

Genbank Accession No. AC079988 teach a human chromosome 2 clone RP11-

795C1, which is isolated from the human BAC library RPCI-11 according to AC079988

(gi:18873965), which contains an insert having 94.6 % sequence identity to instant SEQ

ID NO:100. Note the sequence alignment of instant SEQ ID NO:100 against Genbank

Accession No. AC079988 with gi:10800346 (Exhibit D). Genbank Accession No.

AC079988 sequencing vector M13, and Dye-primer ET, and other means to detect the

nucleic acid inserted in the clone.

As for claims 7, and 8, Voet et al., (1900, Biochemistry, John Wiley & Sons,

pages 839-844) teach that a clone is in a host cell containing a nucleic acid of insert in

an appropriate vector.

As for claims 15, and 17, the intended use in claim 15, and the preamble

recitation in claim 17 are merely suggestive of an intended use and is not given

patentable weight for purposes of comparing the claim with the prior art. The claim

reads on nucleic acids *per se,* and a means *per se.*
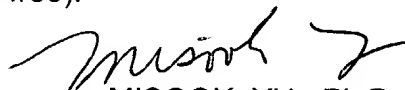
Thus, Genbank Accession No. AC079988 anticipates claims 1, 3-5, 7, 8, 15, and

17.

Conclusion

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to MISOOK YU, Ph.D. whose telephone number is 571-

272-0839. The examiner can normally be reached on 8 A.M. to 5:30 P.M., every other

Friday off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Jeffrey C Siew can be reached on 571-272-0787. The fax phone number

for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free).

MISOOK YU, Ph.D.
Examiner
Art Unit 1642

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:      November 23, 2004, 08:50:59 ; Search time 16.5 Seconds
                                          (without alignments)
                                          4.832 Million cell updates/sec

Title:           us-09-989-920-100
Perfect score:   5074
Sequence:        1 gccagaagcagcccagctt.........aaaaataaaagatgcggcc 2754

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Database :       US09989920.pep:*

Searched:        120 seqs, 14475 residues

Total number of hits satisfying chosen parameters:      240

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=soft -Q=us-09-989-920-100 -DB=US09989920.pep
-SUFFIX=pto -OUT=align100 -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits
-START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -NCPU=6 -NO_XLPXY
-NEG_SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	254.5	5.0	146	1	US-09-989-920-237	Sequence 237, App
2	117	2.3	30	1	US-09-989-920-231	Sequence 231, App
3	109	2.1	213	1	US-09-989-920-206	Sequence 206, App
4	109	2.1	213	1	US-09-989-920-251	Sequence 251, App
5	83.5	1.6	785	1	US-09-989-920-218	Sequence 218, App
6	80	1.6	785	1	US-09-989-920-218	Sequence 218, App
c 7	73	1.4	188	1	US-09-989-920-240	Sequence 240, App
c 8	70.5	1.4	196	1	US-09-989-920-224	Sequence 224, App
c 9	69	1.4	146	1	US-09-989-920-237	Sequence 237, App
c 10	69	1.4	149	1	US-09-989-920-187	Sequence 187, App
c 11	69	1.4	213	1	US-09-989-920-206	Sequence 206, App
c 12	69	1.4	213	1	US-09-989-920-251	Sequence 251, App
13	67	1.3	109	1	US-09-989-920-223	Sequence 223, App
14	66	1.3	456	1	US-09-989-920-175	Sequence 175, App
15	62	1.2	183	1	US-09-989-920-245	Sequence 245, App
16	60.5	1.2	86	1	US-09-989-920-243	Sequence 243, App
17	59	1.2	91	1	US-09-989-920-205	Sequence 205, App
18	59	1.2	451	1	US-09-989-920-210	Sequence 210, App
c 19	58.5	1.2	121	1	US-09-989-920-198	Sequence 198, App
c 20	57.5	1.2	171	1	US-09-989-920-282	Sequence 282, App
c 21	57.5	1.1	196	1	US-09-989-920-224	Sequence 224, App
22	56.5	1.1	162	1	US-09-989-920-199	Sequence 199, App
23	56.5	1.1	456	1	US-09-989-920-175	Sequence 175, App
c 24	55.5	1.1	101	1	US-09-989-920-184	Sequence 184, App
25	55	1.1	594	1	US-09-989-920-200	Sequence 200, App
26	55	1.1	594	1	US-09-989-920-249	Sequence 249, App
c 27	54.5	1.1	224	1	US-09-989-920-274	Sequence 274, App
28	54	1.0	100	1	US-09-989-920-264	Sequence 264, App
c 29	52.5	1.0	240	1	US-09-989-920-247	Sequence 247, App
c 30	52	1.0	130	1	US-09-989-920-208	Sequence 208, App
31	52	1.0	489	1	US-09-989-920-185	Sequence 185, App
c 32	51.5	1.0	610	1	US-09-989-920-212	Sequence 212, App
c 33	50.5	1.0	53	1	US-09-989-920-230	Sequence 230, App
c 34	49	1.0	101	1	US-09-989-920-184	Sequence 184, App
c 35	48	1.0	128	1	US-09-989-920-220	Sequence 220, App
36	48	0.9	252	1	US-09-989-920-204	Sequence 204, App
c 37	48	0.9	1533	1	US-09-989-920-259	Sequence 259, App
38	47.5	0.9	57	1	US-09-989-920-234	Sequence 234, App
39	47.5	0.9	88	1	US-09-989-920-196	Sequence 196, App
40	47.5	0.9	1533	1	US-09-989-920-259	Sequence 259, App
41	47	0.9	149	1	US-09-989-920-279	Sequence 279, App
42	46.5	0.9	149	1	US-09-989-920-187	Sequence 187, App
c 43	46	0.9	107	1	US-09-989-920-233	Sequence 233, App
c 44	46	0.9	183	1	US-09-989-920-245	Sequence 245, App
45	45.5	0.9	63	1	US-09-989-920-193	Sequence 193, App

```
                              ALIGNMENTS

RESULT 1
US-09-989-920-237
; Sequence 237, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 237
;     LENGTH: 146
;     TYPE: PRT
;     ORGANISM: Homo sapien
US-09-989-920-237

Alignment Scores:
Pred. No.:                    1.21e-13        Length:         146
Score:                        254.50          Matches:        59
Percent Similarity:           76.62%          Conservative:   0
Best Local Similarity:        76.62%          Mismatches:     7
Query Match:                  5.02%           Indels:         15
DB:                           1               Gaps:           1

us-09-989-920-100 (1-2754) x US-09-989-920-237 (1-146)

Qy   2139 ATGGGGCCAGGGTCCCACA-GTACTCATGGGCCGAGCCAGGTTGAGCTTCCTGGTGTC 2197
               ||||||||||||||  |  ||||||||||||||  |  ||||||||||||||||||||
Db      1 MetGlyAlaArgValProHisAlaAlaAspGlyProSerGlnValGluLeuProGlyVal 20

Qy   2198 CAGTCCGGATC-CCACTTGCAGATCTCATGCTTCAGATAGGTGGACAAGTTCTTTGT 2256
               |||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db     21 GlnSerGlySerProLeuAlaAspLeuMetLeuSerAspArgTrpAspPhePheCys 40

Qy   2257 CACAGTGCTGGCTC-TGTCCTGAGGCTCATTGCTGGGTGTGCTC--TGCTGGGAA 2313
               |||||||||||  | ||||||||||||||||||||||||||||  ||||||||
Db     41 HisSerAlaGlyLeuCysProGluAlaSerLeuLeuAlaGlyCysAlaHisAlaTrpGlu 60
```

```
Qy  2314 AAG----------------------------CTTTGCGGGGCTTGC 2331
                                         ||||||| ||| |||
Db    61 LysAlaTrpAlaAlaAsnTyrGlyHisThrCysSerLeuCysGlyHisCys 77
```

RESULT 2
US-09-989-920-231
; Sequence 231, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 231
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-231

```
Alignment Scores:
Pred. No.:             0.000438    Length:        30
Score:                 117.00      Matches:       26
Percent Similarity:    96.30%      Conservative:   0
Best Local Similarity: 96.30%      Mismatches:     1
Query Match:           2.31%       Indels:         1
DB:                    1           Gaps:           0
```

us-09-989-920-100 (1-2754) x US-09-989-920-231 (1-30)

```
Qy  2540 ATGGGTCAGGTGGGTGTGGAGAGGTCCTGGGGAGGTGG--GCGTTGAGCTCAAAGTTGTCC 2597
              ||||||||||||||||||||||||||||||||    ||||||||||||||||||
Db     1 MetGlyGlnValGlyValArgGlyProGlyGluValArgAlaLeuSerLysLeuSer 20

Qy  2598 TACTGCCATGTTTTTGTACCT 2618
              |||||||||||||||||||
Db    21 TyrCysHisValPheValPro 27
```

RESULT 3
US-09-989-920-206
; Sequence 206, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 206
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-206

```
Alignment Scores:
Pred. No.:             0.00197     Length:       213
Score:                 109.00      Matches:       73
Percent Similarity:    36.19%      Conservative:  24
Best Local Similarity: 27.24%      Mismatches:    81
```

```
Query Match:           2.15%       Indels:    91
DB:                    1           Gaps:      15
```

us-09-989-920-100 (1-2754) x US-09-989-920-206 (1-213)

```
Qy  1554 GTTGATGCGGTGGAGAAGGACTTTGAGCAGTGGGCAGCAGTGGCCTCTGGCCAGCTC 1613
              ||||| ||||||||||||||||||||||||  ||||:::
Db     9 ValAspGlnThrProArgThrLeuAlaThrMetGlyGlnArgAlaLeuProSerLeu 28

Qy  1614 ACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGACTGA 1673
                            ||||||||| :::||||||||||:::      :::
Db    29 ArgPro-LeuSerProProAlaAlaCysSerGlyAs 45

Qy  1674 GGACCT-TTTGGCTTCTCTGGAGCCTGCAAGCCTCCATGTGTCCAGCTGTCTTCC 1732
          :                             |||             :::
Db    45 pProGlyCysGly--SerGlyAlaGlyLeuPro------------SerAlaSe 58

Qy  1733 TGCTACAAAGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCC 1792
                    |||    |||  ||| :::       |||     |||
Db    58 rAlaAlaAlaGlyIleAlaSerSerAlaValGluAlaValCysGlyAspAlaAlaProAl 78

Qy  1793 GGCCCTCCATAAGGTATCCTGGGCTGAGAATTCTGCATCTGCCATTGGAGGATGGACA 1852
                                          |||       |||||
Db    78 aCysLeu--------------LeuArgThrProLeu----- 85

Qy  1853 GCCTCAAATGGAAGAGTCCACGGAGGAGATGGGTCCGAGTCC------GGCTGTGGC- 1904
          |||::::  ::::      ::    |||||
Db    86 -------------ArgGlyLeuLeuLysProThrGlyProArgSerThrMetGluCysProPr 102

Qy  1905 --------------CATCCAGCCCCTGTGGCTTGTCCAGCCTCGTGTCCACCCCTGTGTC 1951
                        |||||||       :::::
Db   102 oAlaLeuIleValHisProProThrThrGlyMetAlaArgArgAlaAlaSer------- 119

Qy  1952 TTCACTCCAGGGCAGACAGCACCAGCCACTGCAGTTCTTTTTCGTGGAGTAACAGTAGTGA 2011
                                  |||       ||||||:::
Db   120 ------GlnProTrpAlaAlaSerAlaThrProMetLeuSerSerLysAlaSer- 136

Qy  2012 TAGCAGCTGGGGCTAACAGGCTGAGGCTTGGTGTTCTGCCATTTGGTCAGCTTCTCACTC 2071
                                  :::            ||||        :::
Db   137 --------------------LeuCysIle---------ProThrGluAr 143

Qy  2072 GATCCTCCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGG 2131
                            |||           ||| |||
Db   143 gProProGlnProLeuMetArgThrPro--------------------- 153

Qy  2132 AGGTTAGATGGGGGCCAGGGTCCCAACGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCT 2191
                              |||||        |||
Db   154 ------------AlaAlaArg-SerHis----------TrpProIleProHisProAlaSerT 167

Qy  2192 GG---TGTCCAGTCCGATCCCACTT------------------GCCAGATCTCATGC 2227
               |||:::      ||:::|||
Db   167 hrAlaCysProAlaProLeuProValValLeuValAlaProArgSerThrIleLeuSerM 187

Qy  2228 TCTTCAGATAGTGG-------------------GACAAGTTCT 2251
          ::|||||                   ::|||
Db   187 etSerArgThrTrpTh-CysArgArgTrpAlaValAlaAlaProCysArgAlaGluLysLeuM 207

Qy  2252 TTTGTCAGTGCTGGCTCT 2271
          |||  |||:::
Db   207 eCysSerSerArgSer 213
```

RESULT 4
US-09-989-920-251
; Sequence 251, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920

```
;  CURRENT FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 60/252,500
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 284
;  SOFTWARE: PatentIn version 3.1
;  SEQ ID NO 251
;  LENGTH: 213
;  TYPE: PRT
;  ORGANISM: Homo sapien
US-09-989-920-251

Alignment Scores:
Pred. No.:              0.00197      Length:            213
Score:                  109.00       Matches:            73
Percent Similarity:     36.19%       Conservative:       24
Best Local Similarity:  27.24%       Mismatches:         81
Query Match:            2.15%        Indels:             91
DB:                                  Gaps:               15

us-09-989-920-100 (1-2754) x US-09-989-920-251 (1-213)

Qy 1554 GTTGATGCGTGGAGGAGGACTTGAGCAGTGTGGGCAGCAGTGGCCTCCTGGCCAGCTC 1613
              |||||||||||||||||||:::: |||||||:||: |||||
Db    9 ValAspGlnThrProArgThrLeuAlaThrMetGlyGlnArgAlaLeuProSerSerLeu 28

Qy 1614 ACACTCTTGTCCTGGAGGGGCAGCCTAGTCCACCTAGTACCTTGGGACTGA 1673
              ::|||:||||||||||  :: ||
Db   29 AlaLeuLeuSer-------ArgPro-LeuSerProProAlaAlaCysSerGlyAs 45

Qy 1674 GGACCT-TTTGGCTTCTCTGGAGCCTTGCCAAGCCTCTTCCCATGTGTCCAGCTGCTTTCC 1732
              :  |||||| ||||||||  :: |||
Db   45 pProGlyCysGly--SerGlyAlaGlyLeuPro--------SerAlaSe 58

Qy 1733 TGGTACAAAGGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCC 1792
              ||| ::: |||
Db   58 rAlaAlaAlaGlyIleAlaSerAlaValGluAlaValCysGlyAspAlaAlaProAl 78

Qy 1793 GGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACA 1852
              |||
Db   78 aCysLeu----------------------------LeuArgThrProLeu--------- 85

Qy 1853 GCCTCAAATGGAGGAGGCCACGGAGATGGTCCGAGGTCC-----GGCTGTGGC-- 1904
              ||||| ::: ||||
Db   86 --------ArgGlyLeuLeuLysProThrGlyProArgSerThrMetGluCysProPr 102

Qy 1905 --------------CATCCAGCCCCCTGGCCTTGTCCAGCCTCGTGCACCCCTGGTGTC 1951
              |||||||| |||
Db  102 oAlaLeuIleValHisProProThrGlyGlyMetAlaArgAlaAlaSer-------- 119

Qy 1952 TTCACTCCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGGTAACAGTAGTGA 2011
              ||| |||:::|||
Db  120 ------GlnProTrpAlaAlaAlaSerAlaThrProMetLeuSerSerLysAlaSer-- 136

Qy 2012 TAGCAGTGGGCTAACAGGCTAGGCTTTGTTGTTCTGCGCATTGGTCTGCAGCTTCTCACTC 2071
              ||||| :::
Db  137 ------------ProThrGluAr 143

Qy 2072 GATCCTCCTAAAGCAATGGGGAGGGCCCCCACTAGCCAGTTTCAGGAAGTCAACTGGG 2131
              ||||| ::: |||||
Db  143 gProProGlnProLeuMetArgThrPro--------------- 153

Qy 2132 AGGTTAGATGGGGGCCAGGGTCCCAGCTACTGATGCCCAGCCAGGTTGAGCTTCCT 2191
              |||||||| |||||
Db  154 ---------AlaAlaArg-SerHis--------TrpProIleProHisProAlaSerT 167

Qy 2192 GG---TGTCCAGTCCGGATCCCACTT--------- 2227
              |||| |||||
Db  167 hrAlaCysProAlaProLeuProValValLeuValAlaProArgSerThrIleLeuSerM 187

Qy 2228 TCTTCAGATAGGTGG---------- 2251
              ::: ||||
Db  187 etSerArgThrTrpThrCysArgArgTrpAlaValAlaAlaProCysArgAlaLeuLysLeuM 207
```

```
Qy 2252 TTTGTCACAGTGCTGGCTCT 2271
              ||| |:::: |||
Db  207 etCysSerSerArgSer 213

RESULT 5
US-09-989-920-218
;  Sequence 218, Application US/09989920
;  GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto
;  APPLICANT: Recipon, Herve
;  APPLICANT: Chen, Sei-Yu
;  APPLICANT: Sun, Yongming
;  APPLICANT: Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
;  FILE REFERENCE: DEX-0291
;  CURRENT APPLICATION NUMBER: US/09/989,920
;  CURRENT FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 60/252,500
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 284
;  SOFTWARE: PatentIn version 3.1
;  SEQ ID NO 218
;  LENGTH: 785
;  TYPE: PRT
;  ORGANISM: Homo sapien
US-09-989-920-218

Alignment Scores:
Pred. No.:              0.139       Length:            785
Score:                  83.50       Matches:            52
Percent Similarity:     33.16%      Conservative:       11
Best Local Similarity:  27.37%      Mismatches:         63
Query Match:            1.65%       Indels:             64
DB:                                 Gaps:               11

us-09-989-920-100 (1-2754) x US-09-989-920-218 (1-785)

Qy 1463 AGAGCCCTTTGCTTCCTTCTTGCCCGCTGTTCCTGGGGACCCGCTGGCCTT 1522
              :::|||
Db  655 ArgThrLeuValTyrPro------------LeuGluLeuProLysGluPro------ 667

Qy 1523 TGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGCAG 1582
              |||:::
Db  668 ------ThrSerProProPheArgPro-------CysProGluProAsp--GluLys 681

Qy 1583 TGGTGGCCAGCAGTGGCCTCCTGGCCAGCTCACACTCTGTCCTGGGGAGGGGCAGCCTGA 1642
              ||| :::
Db  682 LeuTrpAspProValGlyTyrTyrTyrSerAsp------------------ 692

Qy 1643 TCTCACCTCCACCTAGTACCTTGGGACTGAGGACCTTTGGCTTCTCTGGAGCCTGCAA 1702
              |||
Db  693 GlySerLeuLysIleValProGlyHisAlaArg--------------------- 703

Qy 1703 GCCTCTTCCCATGTGTCCAGCTGCTTCTCGTCACAAAGGGGACTGCTCACAGTGGCCT 1762
              |||
Db  704 ------------------CysGlnProGly-------------------- 707

Qy 1763 CAGCTTGGTGGTTTGAGGGGCCGCCCCCCCCCCGCCCTCCATAAGGGTATCCTGGGCGAG 1822
              ||||||||||||||||
Db  708 ------GlyGlyProSerProProGlyIlePro----GlyGlnProLeuProSer 724

Qy 1823 AATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAATGA-----------AGGAGT 1870
              :::|||
Db  725 ProThrArgLeuHisLeuGlyGlyLyArgAsnSerAsnAlaAsnGlyTyrValArgLeu 744

Qy 1871 CCCACGGGAGGATGGGTCCGAGTCGGCCATCCAGCCCCCTGTGGCTTGTCCAG 1930
              |||||||||||||
Db  745 GlnLeuGlyGlyGluAsp-ArgGlyGlyLeuGlyHisProLeuPro----------Gl 760

Qy 1931 CCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCCACTGCAGTTCCTTT 1990
              :::|||
Db  760 uLeuAlaAspGluLeuArgGlyLysLeuGlnLeuGlnArgGln------------ProLe 775
```

```
Qy  1991 CTTCGTCGAGTAACAGTAGTGATAGCAGC 2018
              ||||||| :::|||||||
Db   775 uProAspSerAsnProGluSerSer 784

RESULT 6
US-09-989-920-218
; Sequence 218, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
;   LENGTH: 785
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-218

Alignment Scores:
Pred. No.:              0.245        Length:      785
Score:                 80.00        Matches:      89
Percent Similarity:    29.56%       Conservative: 31
Best Local Similarity: 21.92%       Mismatches:  119
Query Match:            1.61%       Indels:      169
DB:                     1           Gaps:         18

us-09-989-920-100 (1-2754) x US-09-989-920-218 (1-785)

Qy  2178 TGGCTCGGCCATCAGTAGTGTGGGACCCTGGCCCCATCTAACCTCCCAGTTGACTTC 2119
              |||||||:::::|||||||||||||| :::|||
Db   398 TrpLeuArgLysAlaIleSerLeuGlyProTrp-ValHisLeuIleGluGlnLeu---- 415

Qy  2118 CTGAAAACTGGGCTAGTGGGGGCCTCCCCATTGCTTAGGGAGGATGAGTGAGAAGCTG 2059
              ||||||| :::|||| |||||
Db   416 ------------------GlnLeuPheAspGlnGluProMetArgSerLe 426

Qy  2058 ACCAAATGCGCAGAACACAAGACCAGAAGCCTAGCCTGTTAGCCCCAGCTGCTATCACTACTGTT-- 2001
              ::|||||::::  ||||||| |||||
Db   426 uValLeuSerGlnSerLysValLysLeuLeuPheAlaGlyLeuLeuPheAlaGlySerArgSerGlnLeuValGl 446

Qy  2000 --------ACTCACGAAGAAAGAAGAACTCAGTGGCTGCTGTC------- 1967
              ::: |||||:||:  ||||:
Db   446 nLeuProValAlaAspCysMetLysTyrArgSerCysAlaAspCys-ValLeuAlaAargA 466

Qy  1966 ---TGCCCCTGAGTGAAGACACCAGGGGTGCACAGAGGCTGGACAAGCCACAGG 1915
              ::|||| |||:|:  :::
Db   466 spProTyrCysAlaTrpSerValAsnThrSerArgCysValAlaValGly-GlyHisSer 485

Qy  1914 GGG------CTGGATGGCCACAGCCCGGACCTCGGACCTCCCGTG---GGACTCCTT 1864
              |||||||  :::
Db   486 GlySerLeuLeuIleGlnHisValMetThrSerAspThrSerGlyIleCysAsnLeuArg 505

Qy  1863 CCATTGGAGGCTGTCCATCCTCCA-------------------- 1840
              :::
Db   506 GlySerIlysLysValArgProThrProLysAsnIleThrValAlaAlaGlyThrAspLeu 525

Qy  1839 --------ATGCAGATGCAGAA------TTCTCAGGCCCA 1813
              ::|||||
Db   526 ValLeuProCysHisLeuSerSerAsnLeuAlaHisAlaAArgTrpPheThrPheGlyArg 545

Qy  1812 GGATACCCTTATGGAGGGCGGGG------------- 1789
              ||||||||
Db   546 AspLeuProAlaGluProGlySerPheLeuTyrAspAlaArgProAlaArgLeuGlnAlaLeuVal 565
```

```
Qy  1789 ------------------------------------------- 1789
Db   566 ValMetAlaAlaGlnProArgHisAlaGlyAlaTyrHisCysPheSerGluGluGlnGly 585

Qy         ---------------------------GGCGGCCCCTCAAA-ACCACC 1769
              |||||||
Db   1788 ------------------------------------------- 1788

Db   585 AlaArgLeuAlaAlaGluGlyTyrLeuValAlaValAlaValAlaGlyProSerValThrLeu 605

Qy  1768 AAGCTGAGGCCA--- 1757
              ::: |||
Db   606 GluAlaArgAlaProLeuGluAsnLeuGlyLeuValTrpLeuAlaValaValAlaLeuGluGly 625

Qy  1756 -------------CTGTGAGCAGTCCCCTTTGTAGCAGGAAGAGACAGCTGGACAC 1715
Db   626 AlaValCysLeuValLeuValLeuLeuValLeuSerLeuValLeuArgLeuArgArgGluGlu 645

Qy  1714 ATGGGAAGAGGCTTGCAGGCTCCAGAGAAG------------CCAAAA 1679
              ::: :::|||:  :::|||
Db   646 LeuGluLysGlyAlaLysAlaThrGluArgThrLeuuArgThrLeuVaTyrProLeuGluLeuProLys 665

Qy  1678 GGTCCTCAGTCCCCAAGGTACTAGGTGGAGGTGAGATCAGGCTGCGCCCTCCCAGGACAAG 1619
              |||||||||||||||||:  :::
Db   666 GluProThrSerProProPhe------------ArgProCysProGluProAspGlu 680

Qy  1618 AGTGTGAGCTG------- ...CCA 1604
              ::: :::
Db   681 LysLeu---TrpAspProValGlyTyrTyrTyrSerAspGlySerLeuLysIleValPro 699

Qy  1603 GGAGGCCACTGCTGCCCACCACT-GCTCAAAGTCCTTCTCCACGGCCATCAACCCTGAGGGA 1545
              |||||||||||||||||||  :::
Db   700 GlyHisAlaArgCysGlnProGlyGlyGlyProProSerPro------ProProGly 716

Qy  1544 ---CCTGGCCAGGGGATGCAGACCAAAGGCCCAGGCGGGTCCCAGGAAGCAGGAAGCAGGC 1488
              ::: :::  |||
Db   717 IleProGlyGlnProLeuProSer------------ProThrArgLeuHisLeuGly 731

Qy  1487 GGCAAGAAGGAAAAGCAAAGGGGCTCTCGCATTCCACCACGAGGCGGGGTGGACTCTCGCAGG 1428
              ||||::: |||||:  :::
Db   732 GlyGlyGlyArgAsnSerAsnAlaAsnGlyTyrValArgLeuGlnLeuGlyGlyGluAspArg 751

Qy  1427 AGCCAGCTGAGCACCACCAGAACCTTCCAAGGGGGCT---- 1392
              |||
Db   752 GlyGlyLeuGlyHisPro---LeuProGluLeuAlaAspGluLeuArgArgLysLeuGln 770

Qy  1391 ---CGCCAGCCA 1383
              ||||||
Db   771 GlnArgGlnPro 774

RESULT 7
US-09-989-920-240
; Sequence 240, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 240
;   LENGTH: 188
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-240

Alignment Scores:
Pred. No.:    0.648        Length:    188
```

```
Score:                 73.00          Matches:         51
Percent Similarity:    29.36%         Conservative:    13
Best Local Similarity: 23.39%         Mismatches:      76
Query Match:           1.44%          Indels:          79
DB:                                   Gaps:            14

us-09-989-920-100 (1-2754) x US-09-989-920-240 (1-188)

Qy  1373  GTGTCCTGTGTGGCTGGCGAGCCCCCTTGAAGGTTCTGGTGCTCCTGC  1432
               ||||||||||||
Db    15  ValGlyCysGlyGlyGly----------GlyGlyGlyAlaGlyAlaGlySerTrp  29

Qy  1433  AGAGTCCACCCCGCTCGTGGTGGGAATGCAGAGCCCTTGCTTCCTTCTTGCGCCTG  1492
               |||
Db    30  --ValTrpMetGlyGlyTrp------------------------------  35

Qy  1493  CTTCCGTTCCTGGGACACCCGCTGGGCTTGGTGCTGCATCCCTGGCCAGGTCCCTCAG  1552
               |||
Db    36  --------GlyGlyGlyAlaGlyAlaLeuTrpValAlaValAlGlyAla------  50

Qy  1553  GGTTGATGCGTGGAGAAGGACTTGAGCAGTGGTGGGCAGCAGTGGCTCCTGGGCCAGCT  1612
               ::|||||
Db    51  ----------ArgTrpTrpGlyGlyAlaGlyTrpGlySerCys  61

Qy  1613  CACACTCTTGTCCCGGGAGGGGCCAGCC-------TGATCTCACCTCCACCTAGTACCTTGG  1666
               :::|||||||||  |||
Db    62  GlyArgValLeuValGlyGlyArgAlaValValAlaValGlyValValAlGlyTrp  81

Qy  1667  GGACTGAGGACCTTTTGGCTTCCTCGGAGCCTCCAAGCCTCTTCCCATGTGTCCAGCTGC  1726
               :|| |||
Db    82  Gly--------Trp--------TrpArgValValVal------------  88

Qy  1727  TCTTCCTGCTACAAAGGGGACTGCTCACAGTGG--------CCTCAGCTTGGTGGTTTTGAG  1780
               :: ||||  :::|||||
Db    89  AlaGlyValCysGlyGlyGlyTrpGlyGlyThrArgTrpTrpArgAlaGlyValValGlyGlyGlyGlyGly  108

Qy  1781  GGGCCCCCCCGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGGGCATCTGCCATT  1840
               ||||||
Db   109  GlyAlaValSerGlyPro-----SerGlyAlaGlyProGlyArgArgCysSerMetVal  126

Qy  1841  -----------------GGAGGATGGACAGCCTCAAATGGAGGAGGAGTCCC  1873
               |||||||
Db   127  GluArgArgArgGlyHisValGlySerGlyGlyTrpGlyGlyTrpAlaGlyProGly------Val  144

Qy  1874  ACGGGAGA-TGGGTC--------------------------------CCGAGGTCCGGGCT  1899
               ||  |||
Db   145  ValGlyValTrpAlaArgCysValLeuValAlaGlyAlaValTrpArgArgTrpGlyGlyAla  164

Qy  1900  GTGGCCATCCACGCCCCCTGGCTGTGCCAGCCTCTGTGCACCCGTGTGTCTT  1953
               :||  |||
Db   165  ValTrpGluTrpArgGlyLeuGlyCysGlyAla---------TrpCysVal  178
```

RESULT 8
US-09-989-920-224
; Sequence 224, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 224
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Homo sapien

US-09-989-920-224

```
Alignment Scores:
Pred. No.:             0.975          Length:          196
Score:                 70.50          Matches:         27
Percent Similarity:    35.58%         Conservative:    10
Best Local Similarity: 25.96%         Mismatches:      46
Query Match:           1.42%          Indels:          22
DB:                                   Gaps:            3

us-09-989-920-100 (1-2754) x US-09-989-920-224 (1-196)

Qy  2694  CCCAAGCACAGATCCCACATACAGACTTCTCGTCCGCACTATGGTACAGTAACAAGTGCA  2635
               ||||||  :::||||
Db   104  ProGlnHisProProThrProThrThrProProHisAlaProThr------  121

Qy  2634  AAATATGCTTTATTTCAGTACAAAAACATGGCAGTAGGACAACTTTGAGCTCAACGCCC  2575
                :::|||||||||
Db   122  ----------------------------ProHisThrProProProThrPro  130

Qy  2574  ACCTCCCCAGG-ACCTCTCACCACCTACCACCTCAACCCATCCAAGGGCCACCACCCCGACCAG  2516
               |||||||||||||||||
Db   131  ProArgProProThrHisThrHis-------------ThrProProHisPro  144

Qy  2515  ATACTCCCACACCTTTAGACAAAGAGT-----CACCCAATCTGGAGAAAGTGTGGAGG  2462
               ||||||  :::|||||
Db   145  ProThrProProProLeuProThrProThrProHisProThrProHisSerHisSerThrLeuSer  164

Qy  2461  TTACATCTTTAGAAAGAAATCATTTAAATACATGAACATTAGAGAACACAGTAACCGTG  2402
               ::: |||  :::
Db   165  ProHisHisProHisSerThrThrSerSerLeuProSerThrHisAsnAsnIleThrAsn  184

Qy  2401  CTTCCACCCAGC  2390
               ||||||:::
Db   185  ThrProProAla  188
```

RESULT 9
US-09-989-920-237
; Sequence 237, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 237
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapien

US-09-989-920-237

```
Alignment Scores:
Pred. No.:             1.2            Length:          146
Score:                 69.00          Matches:         23
Percent Similarity:    36.11%         Conservative:    3
Best Local Similarity: 31.94%         Mismatches:      20
Query Match:           1.39%          Indels:          27
DB:                                   Gaps:            3

us-09-989-920-100 (1-2754) x US-09-989-920-237 (1-146)

Qy  2694  CCCAAGCACAGATCCCACATACAGACTTCTCGTCCGCACTATGGTACAGTAACAAGTGCA  2635
               ||||||  :::|||
Db    85  ProProHisProThrHisProAsnThrHisProArgProGlnThr----------  100
```

```
Qy  2634 AATATGTTTATTCAGTACAAAAACATGGCAGTAGGACAACTTTGAGCTCAACGCCC 2575
                                                            |||||
Db   101 ------------------------------------------ProThrPro 103

Qy  2574 ACCTCCCAGGACCTCTCACACCCACCTG-ACCCATCCAAGGGCCACCACCA---CCCCGA 2519
              ||| |||  ||  |||||                             |||||
Db   104 ThrThrProHisProProThrProThrProProProGlnHisProHisProProArg 123

Qy  2518 CAGATACTCCCACCACACTTTAGAAAAGAGTCACCCA 2483
                                  :::||||||
Db   124 -------------------ProProProThrSerThrHisPro 131
```

RESULT 10
US-09-989-920-187
; Sequence 187, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 187
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-187

Alignment Scores:
Pred. No.:   1.21
Score: 69.00
Percent Similarity: 37.04%
Best Local Similarity: 29.63%
Query Match: 1.39%
DB: 1

Length: 149
Matches: 32
Conservative: 8
Mismatches: 44
Indels: 24
Gaps: 4

us-09-989-920-100 (1-2754) x US-09-989-920-187 (1-149)

```
Qy   329 GGAGGTGGCAATCAGAAGAGGGTACAAACATTTCCTACCTGCACCCTGGACCGGTAAGA 270
              |||||||
Db    25 GlyGlyGlyGlnGlyGlnLys--------------PheArgValAspMetProGly-------- 37

Qy   269 TGACACAGACGCATGCAACCCGAGCGGAAGTGCCTGTGTGAAGCCAGTCATGTGGGAGACC 210
              ||||||||||              :::                  |||
Db    38 ------------------SerGlySerAlaPheIleProThrIleAsnAlaIleThr 50

Qy   209 CGCTCTAGAGACCTTGCCCTGCTCTGGGGACATCGCCAGTGCAATTTACACTGTCTAGA 150
              |||:|||||                  ::|||
Db    51 ThrSerGlnAspLeuGln-----TrpMetValGlnProValIleThrSerMetSer 68

Qy   149 AAAGCTAAGGCCACCAGCGTCATCCGTCCAGGAGGTCACACAAGATCAAAAGAGAGGAT 90
              ||||||||||              ||||
Db    69 AsnProTyrProArgSerHisProTyrSerProLeuProGlyLeuAlaSerValAlaGly 88

Qy    89 CATCGAGCT------------------------ACCGAGGGAGCCGTGTTTTCCCCAAATG 54
              ||||                                 ||| ::
Db    89 HisMetAlaLeuProArgProGlyValIleLysThrThrValGlyThrValGlyArgArg 108

Qy    53 CGAAGGGAACAGCAGTCATCTCCA 30
              |||||:::||||
Db   109 ArgArgAspGluGlnLeuSerPro 116
```

RESULT 11
US-09-989-920-206
; Sequence 206, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 206
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-206

Alignment Scores:
Pred. No.:   1.25
Score: 69.00
Percent Similarity: 39.31%
Best Local Similarity: 27.75%
Query Match: 1.39%
DB: 1

Length: 213
Matches: 48
Conservative: 20
Mismatches: 67
Indels: 40
Gaps: 10

us-09-989-920-100 (1-2754) x US-09-989-920-206 (1-213)

```
Qy   658 CCATCAGCTGCCCTGCCACAGGAGCAGAAAACCATGCTCAGAGCTACTTCCTAAGGACA 599
              |||           ||| |||            :: :: ||
Db    13 ProArgThrLeuAlaThrMetGlyGlnArgAlaLeuProSerLeuAlaLeuLeuSer 32

Qy   598 CTGCCTCTAGAAGGCCCTAAAGTGGCA--------GAGCGCGGGATG--GCAGTGCA 551
              |||||                    |||              |||
Db    33 ArgProLeuSer-ProProProAlaCysSerGlyAspProGlyCysSerGlyAla 52

Qy   550 GCCCTGGCCACTAGCACAGCAGTTGCC--CTCAGATCTGTTGTTAGCTGTGACCCATAGATT 494
              |||            :::::: |||  ||:                   :::
Db    53 GlyLeuProSerAlaSerAlaAlaAlaGlyIleAlaSerAlaSerAlaValGluAlaValCys 72

Qy   493 GGGGAGGGAGGGAAAGCC-------AGGAGGAGTCTTTGCAGAAAGTCTGGAAGCTAAGT 440
                                       ||| :::  |||
Db    73 GlyAspAlaAlaProAlaCysLeuLeuArgThrProLeuArgGlyLeuArgGlyLeuLysProThr 92

Qy   439 GGGGGTTTCAGGAGCTTCAGG-TGCCCCTGGGCAGGTCCCAAGGAGGCCTGGGGTGCC 381
              |||                    |||||:::
Db    93 GlyPro--ArgSerThrMetGluCysPro-------------------- 101

Qy   380 TGACCAGGCCTCA-CTGGCCCTCAAGCCCTGTCCCTAAGGGGCGATGGGAAGGAGGTGG 322
              |||  :::               |||||                   ||||||
Db   102 --------ProAlaLeuIleValHisPro---ProThrGlyGlyMetAlaArgArgAla 117

Qy   321 CAATCAGAAGAGGGTACAAACATTTCCTACCTGCAACCCTGACCCGTAAGATGACACAG 262
              |||||           |||
Db   118 AlaSer------------------------GlnProTrpAlaAlaAlaSerAlaThr 128

Qy   261 AGCATGCAACCCGAGCGGAAGTGCCTGTGTGAAGCCAGTCATGTGGGAGACCCGCTCTAG 202
              :::
Db   129 ProMetLeuSerSerLysAlaSerLeuCys--------------------IleProThrGlu 142

Qy   201 AGACCTTGCCCT---GCTCTGGGGACATCGCCCAGTGCA 166
              |||||||                :::||
Db   143 ArgProProProGlnProLeuMetAlaArgThrProAlaAla 155
```

RESULT 12
US-09-989-920-251
; Sequence 251, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming

RESULT 13
US-09-989-920-223
; Sequence 223, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920

APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
FILE REFERENCE: DEX-0291
CURRENT APPLICATION NUMBER: US/09/989,920
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/252,500
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 251
LENGTH: 213
TYPE: PRT
ORGANISM: Homo sapien
US-09-989-920-251

Alignment Scores:
Pred. No.: 1.25
Score: 69.00
Percent Similarity: 39.31%
Best Local Similarity: 27.75%
Query Match: 1.39%
DB:

Length: 213
Matches: 48
Conservative: 20
Mismatches: 67
Indels: 40
Gaps: 10

us-09-989-920-100 (1-2754) x US-09-989-920-251 (1-213)

Qy 658 CCATCAGCTGCCCTGCACAGGAGCAGAAAACCCATGCCTCAGAGCTACTTCCTAAGGACA 599
            |||| :::::: :::||| ||
Db   13 ProArgThrLeuAlaThrMetGlyGlnArgAlaLeuProSerSerLeuAlaLeuLeuSer 32

Qy 598 CTGCCCTCTAGAAGGCCCTAAAGTGCA-------GAGCCGGGATG--GCAGGTGCA 551
            |||||||| :::|||||||
Db   33 ArgProLeuSerProProProAlaCysSerGlyAspProGlyLySerGlyAla 52

Qy 550 GCCCTGGCCACTAGCACAGCAGTTGCC---CTCAGATCTGTTAGCTGTGACCCATAGATT 494
            || :::::::::||| ::: :::
Db   53 GlyLeuProSerAlaSerAlaAlaAlaGlyIleAlaSerSerAlaValGluAlaValCys 72

Qy 493 GGGGGAGGGAGGAAAGCC-----AGGAGGAGCTTTGCAGAAAAGGTCTGGAAGCTAAGT 440
            |||| || :: :::|||
Db   73 GlyAspAlaAlaProAlaCysLeuLeuArgThrProLeuArgGlyLeuLeuLysProThr 92

Qy 439 GGGGGTTTCAGGAGCTCTCAGGG-TGCCCCTGGGCAGAGTCCCAAGAGGAGGCCTGGGGTGCC 381
            |||||||:::: ||||
Db   93 GlyPro---ArgSerThrMetGluCysPro- 101

Qy 380 TGACCAGCGCCTCA-CTGGCCCTCAAGCCCTGTCCTAAGGGGCCATGGGAAGGAAGGAGGTGG 322
            ||||||| ||| :: :::|||||||
Db 102 -------ProAlaLeuIleValHisPro---ProThrGlyMetAlaArgArgAla 117

Qy 321 CAATCAGAAGAGGGTACAAAACATTTCCTACCTGCAACCCTGACCCGGTAAGACCGGTAAGATGACACAG 262
            |||||||||
Db 118 AlaSer---------------GlnProTrpAlaAlaAlaSerAlaThr 128

Qy 261 AGCATGCAACCCGAGCGGCGAAGTGCCTGTGTGAAGCCAGTCATGTGGGAGACCCGCTCTAG 202
            ||||| :::
Db 129 ProMetLeuSerSerLysAlaSerLeuCys- 142

Qy 201 AGACCTTGCCCT---GCTCTGGGACAATCGCCCCAGTGCA 166
            ||||||| :::::||:||
Db 143 ArgProProProGlnProLeuMetArgThrProAlaAla 155 ---IleProThrGlu 142

CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/252,500
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 223
LENGTH: 109
TYPE: PRT
ORGANISM: Homo sapien
US-09-989-920-223

Alignment Scores:
Pred. No.: 1.61
Score: 67.00
Percent Similarity: 28.87%
Best Local Similarity: 23.94%
Query Match: 1.32%
DB:

Length: 109
Matches: 34
Conservative: 7
Mismatches: 41
Indels: 61
Gaps: 8

us-09-989-920-100 (1-2754) x US-09-989-920-223 (1-109)

Qy 2294 TGGGTGTGCTCGCTGCGGGAAAAGCTTTCGCGGGCTTGCTTGGTTAACCACAGAAGAGAAG 2353
            |||::: ||| ||| :::|||
Db   3 TrpLeuSerValGlyGly-------GlyGlyArgGluTrp-----SerGluMetLeu 17

Qy 2354 GGGACTGTTTGGGGTGCCTCCTCTGCAGCCTCCCCGTGCTGGGTGGGAAGCACGTTACTGT 2413
            ||| ||| ||
Db   18 GlyValValTrp-------------------------------------TrpTrpGlyGlyValGly 27

Qy 2414 GTTCTC--------TAATGTTCATGTATTAAAATGATTTCTTTCTAAAGATGTAA 2461
            ||| |||:::
Db   28 ValTrpValGlyValGlyValCysGlyCysVal---- 38

Qy 2462 CCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGTGGGAGTATCTGTCG 2521
                                  ||| ::: :::
Db   39 --------------------TrpTrpValValGlyValValTrpProTrpArgCysValGly 52

Qy 2522 GGGTGGTGTGGCCCTTGGATG--GGTCAGGTGGGTGTGAGAGGTCC-TGGGGAGGTGGG 2577
            |||| |||||||
Db   53 CysGlyCysValValTrpTrpGlyGlyValValGlyValValGlyValGlyCysTrpGlyGly--- 71

Qy 2578 CGTTGAGCTCAAAGTGTCCTACTGCCATGTTTTGTACCTGAAATAAAGCATATTTGC 2637
            |||
Db   71 - 71

Qy 2638 ACTTGTTACTGTACCATA----GTGCGACGAGAAGTCTGTATGTGGGATCGTG 2688
            ||| ||| ||||||
Db   72 ---CysValCysValValGlyValCysValGlyValGlyValValGlyArgVal 90

Qy 2689 CTTGGG 2694
            :::|||
Db   91 ValGly 92

RESULT 14
US-09-989-920-175
; Sequence 175, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 175
; LENGTH: 456
; TYPE: PRT

; ORGANISM: Homo sapien
US-09-989-920-175

Alignment Scores:
Pred. No.:   2.17
Score: 66.00
Percent Similarity: 32.51%
Best Local Similarity: 21.86%
Query Match: 1.30%
DB: 1

us-09-989-920-100 (1-2754) x US-09-989-920-175 (1-456)

	Length:	456
	Matches:	115
	Conservative:	56
	Mismatches:	184
	Indels:	174
	Gaps:	28

Qy   885 CCGGTGTGTGACTGACCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCACCTGGCTGT 944
               ||| || ||||||
Db   3   ----------------------ProArgCysProProProCys----------------- 11

Qy   945 TTATGCCACTGATTTATG------ATAGGGAATATATCTTTGAA-------------- 983
               ::: ||| |||||||
Db   12   ---CysGluArgLeuValAlaAsnValAlaGlyLeuArgPheGluThrArgAlaArgThr 30

Qy   984 ----CCCAATGAAGTGTTTTTCTCCCCATCACAAAAAAAAATTCTTA---- 1028
              ||||:| || :::|||
Db   31   LeuGlyArgPheProAspThrLeuLeuGlyAspArgAlaArgArgGlyArgPheTyrAsp 50

Qy   1029 ------------TTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTG 1073
                   ::: ||| |||
Db   51   AspAlaArgGluTyrPheAspArgHisArgProSerPheAspAlaValLeu---- 69

Qy   1074 TATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTT------GCTCTTACCGCA 1127
                   :::   :::::: || |||
Db   70   --------------TyrTyrTyrGlnSerGlyValArgLeuArgArgProAlaHisValPro 85

Qy   1128 ACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTACCGTCTAGCCCGT 1187
              :::||| ||||||||
Db   86   LeuAspValPheLeuGluGluGluValAlaPheTyrGlyLeuGlyLeuValAlaAlaLeuAlaArg 105

Qy   1188 TACAGAAAAAGTCTGCTACTGAGCCAGACCTCCATTAACCTCCATCCTGTTGGATT 1247
                   :::||| |||||||||
Db   106 LeuArgGluAsp----------------------------GluGlyCysProVal 114

Qy   1248 ATTTAAAGAAAAGCCTCAGACAGTAGGCCTTTTTAAAAGAATAAAAATGACTTGGT-TTG 1306
              |||||| ::::::
Db   115 ProProGluArgProLeuProArgArgAlaPheAlaArgGln--------LeuTrpLeu 131

Qy   1307 CGCTTGGAAGCAGCAGGGGAAGCATTCAGATGAGCG------------GTTCTGCATTA 1351
              |||:::   :::
Db   132 LeuPheGluPheProGluSerGlnAlaLeuSerGlnAlaAlaArgValLeuAlaValSerValLeu 151

Qy   1352 ACCCTGCCCATCACGCATCTCGTCTCCTGTGTGGCTGGCGAGCCCCCCTG------- 1402
              |||::: :::::: :::::: |||||
Db   152 ValIleLeuValSerIleValValPheCysLeuGluThrLeuProAspPheArgAspAsp 171

Qy   1403 ---GAAGGTTCTGGT--GCTTCAGCTGGCTCCTCGCAGAGTCCACCCGCCTCGTGGTGG 1456
              ::|||:::|||| :::
Db   172 ArgAspGlyThrGlyLeuAlaAlaAlaAlaAlaAlaGlyProPheProAla-------- 188

Qy   1457 GAATGCAGACGCCCTTTGCTTCCTTCTTGCTTCCTGTTCCTGGGGACCCGCTG 1516
              :::||| ::::::|||||
Db   189 ----------------ProLeuAsnGlySerSerGlnMetProGlyAsnProPro 201

Qy   1517 GGCCTTGGTCGTCGCATCCCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAAGGACTTT 1576
              ||| ::: :::
Db   202 ArgLeuProPheAsnAspProPhePheValGluThrIleuCysBIle----- 217

Qy   1577 GAGCAGTGGTGGGC-AGCAGTGGCCTCCTGGCCAGCTCACACTCTTTGTCCTGGAGGGGC 1635
              ::: ::: |||||
Db   218 -----CysTrpPheSerPheGluLeuValArgLeuLeuValCysProSerLysAla 235

Qy   1636 AGCCTGATCTCACCTCCACCTAGTACCTTGGGACTGGAGGACCTTTGCCTTCCTGGAG-- 1695
              :::   :::|||
Db   236 IlePhePheLys--------AsnValMetAsnLeuIleAspPheValAlaIleLeu-- 251

Qy   1696 CCTGCAAGCCTCTTCCCATGTGTCCA-GCTGCTCTTCCTGCTCTACAAAGGGACTGCTCAC 1754
               ::: |||
Db   252 ------ProTyrPheValAlaLeuGlyThrGluLeuAlaArgGlnArgGlyValGlyValGln 269

Qy   1755 AGTGGCCTCAGCTTGGTGGTTTGAGG--------------GGCCGCCCCCCG 1793
               :::::||||||
Db   270 GlnAlaMetSerLeuAlaIleLeuArgValIleLeuArgLeuValArgValPheArgIlePhe 289

Qy   1794 GCCCTTCCATAAGGGTATCCTGGGCTGAGAATTCGCATCCTGCCATTGGAGGATGGACAG 1853
                                                             |||||
Db   290 LysLeuSerArgHisSerLysGlyLeuGlnIleLeu----------------GlyGln 303

Qy   1854 CCTCAAATGAAGGAGTCCCACGGGAGATGGGT------------------ 1886
              ::: ||||:::||
Db   304 ---ThrLeuArgAlaSerMetArgGluLeuGlyLeuLeuIlePhePheLeuIleGlyGly 322

Qy   1886 --- 1886

Db   323 ValValLeuPheSerAlaValTyrPheAlaGluValAlaAspArgValAspSerHisPhe 342

Qy   1887 ----------CCGAGGTCC---------------------------------GGCGTGGC 1904
                        |||
Db   343 ThrSerIleProGluSerPheTrpTrpAlaValValThrMetThrThrValGlyTyrGly 362

Qy   1905 CATCCAGCCCCCTGTGGC--------------------TTGTCCAGCCCCTGTGTCACCCCTGGTG 1949
              ||||| :::
Db   363 AspMetAlaProValThrValGlyLysLysIleValGlySerLeuCysAla---- 379

Qy   1950 TCTTCACTCCAGGGG----CAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGT 2006
              ::: ||| :::
Db   380 ------IleAlaGlyValLeuThrIleSerLeuProValProValIleValSerArgAsnPhe 397

Qy   2007 AGT----------------GATAGCAGCTGGGGCTAACAGCTAGGCTTTGTGTTCTGCGCCATT 2054
              ||| :::   :::
Db   398 SerTyrPheTyrHisArgGluThrGlyGluGlyGluAlaGlyMet------- 412

Qy   2055 TGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGAGGCCCCCACTAGCCCAGTTT 2114
                                             :::
Db   413 -------------------ValAspMetGlnProCysGlyProLeu 424

Qy   2115 TCAGGAAGTCAACTGGGAGGTTAGAGTGGGGGCCCAGGTCCCACAGCTA----------CTG 2165
              |||||||| :::|||:::|||
Db   425 GluGlyLysAlaAsnGlyGlyLeuValAspGlyLeuValProGluLeuProProLeu 444

Qy   2166 ATGGCCCGAGCCAGT 2181
              |||
Db   445 -TrpAlaProProGly 449

RESULT 15
US-09-989-920-245
; Sequence 245, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; SOFTWARE: PatentIn version 3.1
; NUMBER OF SEQ ID NOS: 284
; SEQ ID NO 245
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-245

Alignment Scores: 3.77
Pred. No.:   183   Length:

Score: 62.00   Matches: 51
Percent Similarity: 29.15%   Conservative: 14
Best Local Similarity: 22.87%   Mismatches: 61
Query Match: 1.22%   Indels: 97
DB: Gaps: 14

us-09-989-920-100 (1-2754) x US-09-989-920-245 (1-183)

Qy   1396 CCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCGAGAGT---CCACCCCGCCTC--- 1449
               ||| || |||| ||||||| |||||||
Db   3 ProAlaGlyValProTrpCysHisLeuGlySerLeuGlnProLeuProProArgPheLys 22

Qy   1450 ---GTGGTGGGAATGCAGAGCCCTTGCTTCCTTCTTGCCGCCGCTGCTTCCTTCCTTCCTGG 1506
                 ::::||| ||| ||| :::
Db   23 AlaValPheSerArgLeuAlaProSerLeuGluTyr------------------AlaTrp 36

Qy   1507 GGACCCCGCTGGGCCT--- 1521
               |||
Db   37 AspTyrArgAlaProThrSerHisAlaArgLeuIleSerLeuAlaPheLeuValGluThr 56

Qy   1522 -------------------------------TTGGTCTGCATCCCCTGGCCAGGT------ 1545
                                              ||||||
Db   57 GlyPheSerProThrValAlaArgLeuValSerAsnSerTrpProProValValArgPro 76

Qy   1546 ------CCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGC 1599
                   ||||| |||:::|||
Db   77 ProLeuProSer----------------------------GlnSerAlaGly 84

Qy   1600 CTCCTGGCCAGCTCACACTCTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGT 1659
               ||| ||| ||||||
Db   85 IleThrGlyValGlyProProCysLeuAlaArgProIleLeu--------ProProHis 101

Qy   1660 ACCTTGGGACTGAGGACCTTTTGGCTTCTTCTGGAAGCCTCTTCCCATGTGTC 1719
               ||||| ||: :: |||||| ||||||
Db   102 -Pro--------------PhePhePhePheAspMetGlu-----SerHisAlaIl 114

Qy   1720 CAGCTGCTCTTCCTGCTCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGA 1779
               ||||| ::: : |||||
Db   114 eThrGlnAlaGlyVal-----------------------GlnTrpArgHisLeuGlySerLeuGl 128

Qy   1780 GGGGGCCGCCCCCC-- 1792
               : ||||||||
Db   128 nProProProMetPheLysAlaSerCysSerLeuLeuSerSerTrpAspTy 148

Qy   1793 ----GGCCCCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATG 1848
                   |||||| |||
Db   148 rArgArgProPro------------------ProArgPro-----AlaIlePheCysIl 160

Qy   1849 GACAGAGCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATC 1908
               ::: :: ||||| ||||| |||
Db   160 ePheSerArgAspGlyValSerProCysAlaProGly------------TrpSerAr 175

Qy   1909 CAGCCCC 1915
               |||||||
Db   175 gSerPro 177

RESULT 16
US-09-989-920-243
; Sequence 243, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prob
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 243
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-243

Alignment Scores:
Pred. No.:   4.44   Length: 86
Score: 60.50   Matches: 21
Percent Similarity: 33.33%   Conservative: 9
Best Local Similarity: 23.33%   Mismatches: 19
Query Match: 1.19%   Indels: 41
DB: Gaps: 5

us-09-989-920-100 (1-2754) x US-09-989-920-243 (1-86)

Qy   2321 GCGGGGCTTGCTTGGTTAACCACCAGAAGAAGAGGGACTGTTTGGGGTGCCTCTGCAG 2380
               |||||:::||| |||||||||
Db   9 AlaGlyValGlnTrpCysAsnPro-----------------GlySerLeuGln 20

Qy   2381 CCTCCCCGCGTCGGTGGAAGCACGGTTACTGTGTTCTCAATGTTCATGATATTAAAAT 2440
               ::::||||| . ||||| :::||| ||::: ::
Db   21 ProProProGlyPheLysArgPhePheCysLeu---------------CysLeuProSer 36

Qy   2441 GATTCTTCTAAAGATGTAACCTCCACAC-------------------------------- 2470
               ::: ||| ::: |||||
Db   37 SerTrpGlyTyrArgHisThrProProArgProAlaAsnPheCysValPheGlyArgAsp 56

Qy   2471 ---CTTCTCCAGATTGGTGACTCTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGG 2527
               :: |||||| |||||||
Db   57 GlyValSerProCysTrp-------------------------------ProGlyTrp 65

Qy   2528 TGTGGCCCT 2536
               |||||
Db   66 SerLeuSerLeuAspValIleCysAspPro 75

RESULT 17
US-09-989-920-205
; Sequence 205, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 205
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-205

Alignment Scores:
Pred. No.:   5.68   Length: 91
Score: 59.00   Matches: 21
Percent Similarity: 33.80%   Conservative: 3
Best Local Similarity: 29.58%   Mismatches: 19
Query Match: 1.16%   Indels: 28
DB: Gaps: 4

us-09-989-920-100 (1-2754) x US-09-989-920-205 (1-91)

Qy   1353 CCCTGC--------CTATCACGCATCTCGTGTCCTGTCGTGGCTGGCGAGCCCC------ 1397
               |||||| :::||||||:::|||
Db   28 ProCysGlyArgGlyLeuAspLeuLysGlnCysProLeuLeuLeuLeuProTrpLeu 47

Qy   1398 --------------------------------CCTTGGAAGGTTCTGCTTCAGCTG 1424
                                           ||||| |||
Db   48 ThrGlyPheLeuAspHisValHisPheThrGlyProTrpAsp--LeuHisLeuAla 66

Qy   1425 GCTCCTGCGAGAGTCACCCCGCCTCGTCGTGGTGGGAATGCAGAGCGCCCTTTGCTTTCCTTCTT 1484
           :||:||||||| |||
Db   67 SerProAlaGlyLeuIleProAlaArg------------------------------- 75

Qy   1485 GCGCCTGCTTCCTGTTCCTGGGGACCCGCTGG 1517
           |||||||||| ::|
Db   76 AlaProSerPheLeuLeuMetValPheArgTrp 86

RESULT 18
US-09-989-920-210
; Sequence 210, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 210
;   LENGTH: 451
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-210

Alignment Scores:
Pred. No.:   6.48   Length: 451
Score: 59.00   Matches: 28
Percent Similarity: 45.98%   Conservative: 12
Best Local Similarity: 32.18%   Mismatches: 34
Query Match: 1.16%   Indels: 14
DB: 1   Gaps: 3

us-05-989-920-100 (1-2754) x US-09-989-920-210 (1-451)

Qy   2037 CTTTGTTCTGCGCATTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGG 2096
           ||||||||||:|||||::||||::|||||
Db   164 LeuLysValAlaArgLeuTrpIleAlaAsnTyrSer--LeuProArgAlaMetLysArg 182

Qy   2097 CCCCCACTAGCCAGTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCA 2156
           ||||:::
Db   183 LeuGluLysGluIlePro 193

Qy   2157 CAGCT-ACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCGATCCCACTTG 2215
           ||| || ||:||| ::|||
Db   194 GluThrThrArgThrSerGlnMetGlnGluHis------LysSerLeuArgSerLeu 211

Qy   2216 CAGATCTCATGCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCCTGTCC 2275
           |||||||||| ::|||||
Db   212 AsnAsnPheCysSerGlnIleGlyAspAspArgPro-IleSerTyrCysHisPheSerPr 231

Qy   2276 TGAGGGCCTCATTGCTGGCT 2294
           :|: ||:|||||
Db   231 oAsnSerLysMetLeuAla 237

RESULT 19
US-09-989-920-198
; Sequence 198, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming

;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 198
;   LENGTH: 121
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-198

Alignment Scores:
Pred. No.:   6.32   Length: 121
Score: 58.50   Matches: 18
Percent Similarity: 54.05%   Conservative: 2
Best Local Similarity: 48.65%   Mismatches: 12
Query Match: 1.18%   Indels: 5
DB: 1   Gaps: 3

us-09-989-920-100 (1-2754) x US-09-989-920-198 (1-121)

Qy   211 CCCGCTCTAGAGACCTTGCCCTGCCTCTGGGGACATC---GCCCAGTGCAATTTACACTGT 155
           ||| |||||| ||||||
Db   11 ProGlyLeuGlu--------CysSerGlyValIleSerAlaHisCysAsnPheHisLeu 27

Qy   154 CTAGAAAAGCTAAGGCCACCAGCGTCA--TCCACTCTCCAGGAGGTCACA 107
           |||||| |||:|||
Db   28 LeuGlyLysSerSerSerProAlaSerGlnValAlaGluIleThr 44

RESULT 20
US-09-989-920-282
; Sequence 282, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 282
;   LENGTH: 171
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-282

Alignment Scores:
Pred. No.:   7.64   Length: 171
Score: 57.50   Matches: 33
Percent Similarity: 34.75%   Conservative: 16
Best Local Similarity: 23.40%   Mismatches: 58
Query Match: 1.16%   Indels: 34
DB: 1   Gaps: 7

us-09-989-920-100 (1-2754) x US-09-989-920-282 (1-171)

Qy   1627 CAGGACAAGAGTGTGAGCTGCCAGGAGGCCACTGCTGCCACCACTGCTCAAAGTCCTT 1568
           ||| |||||::: |||
Db   31 GlnGlnLysThrLysGlyTrpGly------------Cys-HisThrCysGly------- 43

Qy   1567 CTCCACGCATCAACCCTGACCAGGGATCCGGGGATGCAGAC------CAAAGGCC 1517
           ||||||| ::|||||
Db   44 ----------ProLysAlaGlyPheProGlyGlyHisLeuValLeuSerArgPr 59

Qy   1516 CAGCGGGTCCCCA-----------------------GGAACAGGAAGCAGGCCGGCAAGAGGAAA 1475
                    ||||||
Db   59 oHisAsnSerProProLysTyrArgGlyArgThrArgThrGlnHisThrLy 79

Qy   1474 GCAAAGGGCTCTGCATTCCACCACGAGGCG-------------------GGGTGGACTCT 1433
                 |::: ||||||| :::::||||
Db   79 sArgHisAsnThrGlnAsnHisHisThrAlaThrProAlaHisArgArgGlnArgThrAr 99

Qy   1432 GCAGGAGCCAGCTGAAGCACCAGACCTTCCAGGGGGGCTCGCCAGCCACCAGGGACAC 1373
                 ||| ::: ||| :::
Db   99 gArgGluGlnLysGlyGlnLysGlyGlnLysLysAlaSerSerThrIleThrGlnSe 119

Qy   1372 G-----------AGATGCGTGATAGGCAGGGTTAATGCAGAAACCGCTCAT--------- 1333
                 :   :: |||::: ::::::
Db   119 rHisAspLysArgArgThrMetThrLysThrSerSerThrArgHisArgGlnAs 139

Qy   1332 -CTGAATGCTTCCCCTGCTTCCAGCGCAAACCAAGTCATTTATTCTTTTAAAAAAGCC 1274
                 ::: :::   :::::::|||||
Db   139 pLysSerLysAspArgThrArgGlnLysThrArgAspGluThrArgLysLysPr 159

Qy   1273 C 1273
                |
Db   159 o 159

RESULT 21
US-09-989-920-224
; Sequence 224, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 224
;   LENGTH: 196
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-224

Alignment Scores:
Pred. No.:   7.73   Length: 196
Score: 57.50   Matches: 40
Percent Similarity: 24.74%   Conservative: 7
Best Local Similarity: 21.05%   Mismatches: 66
Query Match: 1.13%   Indels: 78
DB: Gaps: 7

us-09-989-920-100 (1-2754) x US-09-989-920-224 (1-196)

Qy   386 CCCAGGAGCCGCCTCCTTGGGACCTGCCCAGGGGCACCCTGGGAGCTCCTGAAACCCCC------ 439
                |||||||:::|||| |||::||||
Db   72 ProAlaProProHisThrProAlaProProThrArgProThrProProProPro 91

Qy   440 ------ACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCGGCTTTCCTCCGTCCCCCC 493
                ||| ::: |||:::||||
Db   92 ThrHisThrProThrHisThrHis--ProProGlnHisProProProThrPro 110

Qy   494 AATCTATGGGTCACAGCTAACAGATC-TGAGGGCAACTGCTGCTAGTGCCAGGGCTG 552
                |||||| |||
Db   111 ThrThrThrProProHisHisAla----------------ProThrPro 122

Qy   553 CACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGCAGTGTCCTTAGGAAGT 612
                ||| ||| |||
Db   123 HisThrProProThrProThrPro---------------------------------- 130

RESULT 22
US-09-989-920-199
; Sequence 199, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 199
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-199

Alignment Scores:
Pred. No.:   8.9   Length: 162
Score: 56.50   Matches: 38
Percent Similarity: 40.46%   Conservative: 15
Best Local Similarity: 29.01%   Mismatches: 37
Query Match: 1.11%   Indels: 42
DB: Gaps: 7

us-09-989-920-100 (1-2754) x US-09-989-920-199 (1-162)

Qy   1377 CCTGTGGTGGCTGGCAGCCCCCTTGGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAG 1436
                |||:||||||::;|||||
Db   17 ProHisTrpValAlaSer-------------------------------------- 22

Qy   1437 TCCACCCCGGCTCGTGGTGGAATGCAGAGCCCTTTGCTTCCTTCTTGCGCCTGCTTC 1496
                ||| ::: :||||
Db   23 SerAlaAlaProHis-LeuSerLeuPheSerProLysLeuValPheLeu-ThrIleIleV 42

Qy   1497 CTGTTCCTGGGGACCCGCTGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTT 1556
                ||| :::|||
Db   42 alaValGlyGly-------------------------GlyGlnMetLeuLys---- 50

Qy   1557 GATGCGTCGGAGAAGGACTTTGAGCAGTCAGAGCGTGGGCAGCAGTGGCC----------TCCT 1604
                |||:::||||
Db   51 -----ValGluAlaAspLeuGluLysGluThrHisGlyValThrValAlaLysAspSerT 69

Qy   613 AGTCTCTGAGGCATGGGTTTTCTGCTCCTCGTGCAGGGCAGCTGATGGGATAAGGTGGGGAA 672
                |||
Db   130 -- 130

Qy   673 GGACGGTCAGTGCTTGGGCCCCCAGCCTCGCTCCAGCCTGGCGATGGGAGAAACCAAACCATGTC 732
                |||:: ::
Db   131 ----------------------------------ProArgProThrHisThrHisThrProProHisPro 144

Qy   733 CCCAGCGAAGGGCCAGAGAGTGGGAACCTGTCCTCAT------------GCCCTCGT 777
                ||| ||||||
Db   145 ProThrProProProLeuProThrProProHisProThrSerHisSerThrLeuSer 164

Qy   778 CCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTCAGGCCTTCATCAAAG 837
                ||| ::: |||
Db   165 ProHisHisPro---His 169

Qy   838 AGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCAGTTACCGGTGTGTGACT 897
                ||||||||||| :::|||
Db   170 SerThrThrSerSerLeuPro------------SerThrHisAsnAsnIleThr 183

Qy   898 GCCCTTGTCAGCTAGCATACGGTGGGCCCA 927
                ||||||
Db   184 AsnThrProProAlaHisThrLeuThrPro 193

```
Qy    1605 GGCCAGCTCACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCACCTAGTACCTT 1664
            |:  ||::::::         |||          ||||
Db      69 rpLysArgAsnSerIleThrSerSerLeuAlaThrThrArgHisProArg------ProT 87

Qy    1665 GGGG-ACTGAGGACCTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTC---CCATGTGTCC 1720
            ||  | ::::::     ||||        ::|||||||           |||||||||
Db      87 rpHisSerGlnArgLeuCysAlaValAlaLysProLeuAsnLeuPheTrpProCysVal- 106

Qy    1721 AGCTGCTCTTCCTGCTACAAAGGGGACTG 1749
            |||||||          |||||||| |||
Db     107 ------------------LeuGlnArgSerLeu 111

RESULT 23
US-09-989-920-175
; Sequence 175, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE:  2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 175
;   LENGTH: 456
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-175

Alignment Scores:
Pred. No.:              9.44            Length:              456
Score:                  56.50           Matches:             42
Percent Similarity:     37.66%          Conservative:        16
Best Local Similarity:  27.27%          Mismatches:          67
Query Match:            1.14%           Indels:              30
DB:                     1               Gaps:                7

us-09-989-920-100 (1-2754) x US-09-989-920-175 (1-456)

Qy     849 GGATGTTGTTCTCTTTGATGAAGGCCTGAAAACTTCCCCTGGCCCCTGGCTGCCCACCTCA 790
            ||||||||                                |  |:||||
Db      10 GlyCysCysGluArg---------------------------LeuValLeuAsnValAla 20

Qy     789 GGGCTC---CTCAGCAGAAGGACGATGAGACAGTTCCCACTCTGGCCTTCGCTGGGG 733
            ||||||      ||  ::|                     |||||||||
Db      21 GlyLeuArgPheGluThrArgAlaArgThrLeuGlyArgPheProAspThrLeuLeuGly 40

Qy     732 GACACTGGTT----------TGGTTCCCCATCCGCCAGCGCTGGCCAGCTGGGCCCAAG 685
            ||||                 ::::::                       :::
Db      41 AspProAlaArgArgGlyArgPheTyrAspAspAlaArgArgGluTyrPhePheAspArg 60

Qy     684 CACTGACCGGTCCTTC--CCCACCTTATCCCATCA-GCTGCCCTGCACAGGAGCAGAAAA 629
            ||||||||           :::::          ::    ::
Db      61 HisArgProSerPheAspAlaValLeuTyrTyrGlnSerGlyGlyArgLeuArgArg 80

Qy     628 CCC------ATGCCTCAGAGCTACTTCCTAAGGACACTGCCCTCTAGAAGGCCCTAAAGTG 575
            :::|||                                        :::|||
Db      81 ProAlaHisValProLeuAspValPheLeuGlu------------------GluVal 93

Qy     574 GCCAGAGCCGGGGATGGCCAGGTCAGCCCACTAGCACAGCAGTTGCCCTCAGATCT 515
            |||||||                      ||||||||||
Db      94 AlaPheTyrGlyLeuGlyValAlaLeuAla----------ArgLeuArgGlu 108

Qy     514 GTTAGCTGTGACCCATAGATTGGGGAGGGAAAGCCAGGAGGAGAGAGCTTTTGCAGAAAG 455
            ||||||:::                         |||||||  |::
Db     109 AspGluGlyCysProValProProGluProLeuProArgProArgArgAlaPheAlaArgGln 128
```

```
Qy     454 GTCTCGAAGCTAAGTGGGGTTTCAGGAGCTCTCAGGGTGCC 413
            :::||||    |||                   ||||||||
Db     129 LeuTrpLeuLeuPheGluPheProGluSerGlnAlaAla 142

RESULT 24
US-09-989-920-184
; Sequence 184, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE:  2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 184
;   LENGTH: 101
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-184

Alignment Scores:
Pred. No.:              10              Length:              101
Score:                  55.50          Matches:             21
Percent Similarity:     38.67%         Conservative:        8
Best Local Similarity:  28.00%         Mismatches:          30
Query Match:            1.09%          Indels:              16
DB:                     1              Gaps:                3

us-09-989-920-100 (1-2754) x US-09-989-920-184 (1-101)

Qy    1467 CCCTTGCTTGCTTCCTTCCTTGCCGCCCTGCTTCCTCGTTCCTGGGACCCGCTGGGCCTTT--- 1523
            ||||||              |||||       |:::|||:::
Db       9 ProArgProPheLeuTyrHisGlyCysTrpValThrSerGlySerHisHisLeuPhePro 28

Qy    1524 GGTCTGCATCCCCTGGCCAGTCCCCTCAGGTTGATGCGTGGAGAAGGACTTTGAGCAGT 1583
            |||                    ||:::||:::||
Db      29 SerLeuPheProIleSerGln--------------------MetTrpGlyHisGlyLeuAspAsp 43

Qy    1584 GGT------------------------GGGACAGTGGCCTCTGGCCAGCTCACA 1616
            |||                         :::  ||||||||||
Db      44 GlyLeuHisArgSerPheHisLeuCysGluSerLys-SerGlyGlnSerAlaArgThrHi 63

Qy    1617 CTCTTGTCTGGGAGGGGCAGCCTGATCTCACCTCACCTAGT 1659
            |||||||                   |||||||||
Db      63 sLeuCysProGlySerAlaProGlnAsnGlnProProAlaSer 77

RESULT 25
US-09-989-920-200
; Sequence 200, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE:  2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 200
;   LENGTH: 594
```

; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-200

Alignment Scores:
Pred. No.:   11.7   Length: 594
Score: 55.00   Matches: 47
Percent Similarity: 35.53%   Conservative: 23
Best Local Similarity: 23.86%   Mismatches: 69
Query Match: 1.08%   Indels: 59
DB: Gaps: 10

us-09-989-920-100 (1-2754) x US-09-989-920-200 (1-594)

```
Qy 1433 AGAGTCCACCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTCCTTCTTGCGCCTG 1492
             ||| ||||||   |||
Db  268 ------ArgPheHisProGluLeuTrp------------------------CysSerGly 277

Qy 1493 CTTCCTGTTCCTGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAG 1552
                                              ||||||:::::
Db  278 ArgSerValProLeuAspArgGlnGlyTyr---------------GlyGlnIleLysVal 292

Qy 1553 GG-TTGATGCGTGGAGAAGGACTTTGAGCAGTGTGGGCAGCAGTGGCCTC--------- 1602
            |||:|:      ::: :::
Db  293 ValArgAlaAspGlyAspThrLeuSerCysIleCysGlyLysThrLysLeuGlyGluAsp 312

Qy 1603 ---CTGGCCAGCTCACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGT 1659
            |||  |||||| |||    ::|||||
Db  313 MetLeuCysLeuLeuHis---------GlyArgAsnSerMet--AlaProProCysGly 328

Qy 1660 ACCTTGGGACTGAGGACCTTTGGCTTCTCTGGAGCCTCTTCCCATGTGTC 1719
            :::    ||:::||||||      ::: :::
Db  329 AspMet------GluAsnLeuLeuCysAlaThrAsp------SerLeuTyr--------- 341

Qy 1720 CAGCTGCTCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCCAGCTTGGTGGTTTGA 1779
             |||:::   ||||||||
Db  342 ----------------------------LeuAspThrMetGlnValMetLysTrpPheGln 352

Qy 1780 GGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCAT 1839
             ||| ||||||  :::::|||
Db  353 ThrAlaLeuThrArgAlaTrpLysGlyIleAlaHisLysTyrGluPheAspLeuAlaPhe 372

Qy 1840 TGGAGGATGGAC---------AGCCTCAAATGGAAGGAGTCCACGGGAGATGGGTCCGA 1890
             :::::|||           |||||||||
Db  373 GlyGlnLeuAspSerProGlySerLeuLysIleLysLysPheArgSerGlyLysPheMetPro 392

Qy 1891 GGTCCGGCTGTGGCCATC---------------------------------- 1908
             ::: :::
Db  393 PheAsnLeuIleProValIleGlnCysAspAspSerAspLeuTyrPheValSerHisLeu 412

Qy 1909 ------CAGCCCCCGTGGCTTGTCCAGCCTCGTGCACCCCTGGTGTCTT 1953
                  :::||||
Db  413 ProArgGluProSerGluGlyThrProAlaSerSerThrAspTrpLeuLeu 429
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-274

Alignment Scores:
Pred. No.:                  12.4          Length:            224
Score:                      54.50         Matches:            40
Percent Similarity:         33.96%        Conservative:       14
Best Local Similarity:      25.16%        Mismatches:         56
Query Match:                1.10%         Indels:             50
DB:                                       Gaps:                8

us-09-989-920-100 (1-2754)  x  US-09-989-920-274 (1-224)

Qy  1960  TGGAGTGAAGAACACCAGGGGTGCACAGAGGGCTGGACAAGCCACAGGGGGCTGGATGGCCA  1901
                ||||||  |||     ||||                        ||
Db    94  TrpSerLeuAspValAsnGly---------------------------GlyAspGluTrpLys  105

Qy  1900  CAGCCGGACCTCGGACCCATCTCCGTGGGACTCCTTCCATTGAGGCTGTCCATCCTCC  1841
                ||||||                                       ||  |||:
Db   106  ValGluAspLeuSerArgAspGlnArgLysGlu-----------------PheProAsnAsp  120

Qy  1840  AATGCAGATGCAGGAATT---CCTCAGGCCCAGGATACCCTTATGGAGGGCCGGGGGGCGG  1784
                   ||||:     ||          :::|:||
Db   121  GlnValArgSerGlnAlaArgLeuArgValGlnValProAlaValArg-------------  136

Qy  1783  CCCCTCAAAACCACCAAGCTGAGGCCACTGTGAGCAGTCCCTTGTAGCAGGAGAGCA  1724
                 :::                                   |||||
Db   137  ------------------------------SerAlaProValAlaArgAlaArgAla  145

Qy  1723  GCTGGACACATGGGAAGAGGCTTGCAGCCTCCAGAGAGCCA--------AAAGGTCCT  1673
                |||||||  :::                             ||  |||
Db   146  SerGlyAspLeu---------------ProAlaArgProGlyAspHisProAlaGlu  159

Qy  1672  CAGTCCCC-AAGGTACTAGGTGGAGGTGAGATCAGGCTGCGCCCCCCTCCCAGGACAAGAGTGT  1614
               :::   |||||                                     ||||||
Db   160  GluArgCysGlnValGlyLeuProHisGlyLeuProHisIleLeuGlnLeuPro-----------  175

Qy  1613  GAGCTGGCCAGGAGGCCACTGCTGCCCACCACTGCTCAAAGTCCTTCTCCACGC-------  1560
                ||||||||||                     ::::||
Db   176  -----AlaArgArgProLeuHisLeuHisLeuValSerAlaSerArgArgArgGlyHisSerLeuLeu  193

Qy  1559  -------------ATCAACCCTGAGGACCTGCCAGGGGATGCAGACCAAAGGCCCAGCG  1512
                           |||||||         :::      ||||||
Db   194  GlyArgLeuValArgProGluGlyHisGlnGlyHisGlnHisHisHisHisArgAlaProAla  212

RESULT 28
US-09-989-920-264
; Sequence 264, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 264
;   LENGTH: 100
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-264

Alignment Scores:
Pred. No.:                  12.7          Length:            100
Score:                      54.00         Matches:            12
Percent Similarity:         41.03%        Conservative:        4
```

```
Best Local Similarity:      30.77%        Mismatches:         13
Query Match:                 1.06%        Indels:             10
DB:                                       Gaps:                1

us-09-989-920-100 (1-2754)  x  US-09-989-920-264 (1-100)

Qy  1679  TTTGGCTTCTCTGGAGCTGCGCAAGCCTCCTCCATGTGTCCAGCTGCTCCTTCCTGCTAC  1738
                ||:::|||     ||||||||
Db     1  PhePheLeuArgTrpSerLeuAlaGlnValAlaGlnAlaAla------------------  14

Qy  1739  AAAGGGGACTGCTCACAGTGCCTCAGCTTGGTGGTTTGAGGGGTTGAGGGCGCCCCGGC  1795
                                          :::       ||||||
Db    15  ---------------------ArgGlnTrpLeuAsnLeuSerLeuGlnProProProGly  29

RESULT 29
US-09-989-920-247
; Sequence 247 Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 247
;   LENGTH: 240
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-247

Alignment Scores:
Pred. No.:                  16.9          Length:            240
Score:                      52.50         Matches:            44
Percent Similarity:         35.68%        Conservative:       22
Best Local Similarity:      23.78%        Mismatches:         66
Query Match:                 1.03%        Indels:             54
DB:                                       Gaps:                9

us-09-989-920-100 (1-2754)  x  US-09-989-920-247 (1-240)

Qy  1932  CTCTGTGCACCCCTGGTGTCTTCACTCCAGGGCAGCAGCAGCCACTGCAGTCCTTC  1991
                :::|||||| |||:||           ||   :::
Db    21  ValCysAlaThrLeuIleLeuLeuLeuLeuLeuGlyLeuSerGlyLeuGlyLeuSerPhe  40

Qy  1992  TTCGTGAGTAAC---------------------------------AGTAGTGAT------  2012
                ::::::::|||                             |||
Db    41  LeuLeuThrHisArgThrGlyLeuArgSerProAspIleProGlnAspTrpValSerPhe  60

Qy  2013  ---AGCAGCTGGGCTAACAGGCTAGGCTTTGT-----------------GTTCTGCGCATT  2054
                 |||||||||||                                   |||
Db    61  LeuArgSerPheGly---GlnLeuThrLeuCysProArgAsnGlyThrValThrGlyLys  79

Qy  2055  TGGTCAGCTTCTCACTCGATC----------------CTCCCTAAAGCAATGGGG  2093
                |||  :::                          ||
Db    80  TrpArgGlySerHisValValGlyLeuLeuThrThrLeuAsnPheGlyGlyProAsp  99

Qy  2094  AGGCCCCACTAGCCCAGTTTCAGGAGTCAACTGGGAGGTTAGATGGGGGCCAGG-----  2150
                |||                 :::      ||||||||||
Db   100  ArgAsnLysThrArgThrPheGlnAlaThrValLeuGlySerGlnMetGlyLeuLeuLysGly  119

Qy  2151  ------GTCCCACAGCTACTGATGGCCCAGCCCAGGTGAG----------------  2185
                     |||||:::       |||||||
Db   120  SerSerAlaGlyGlnLeuValLeuIleThrAlaArgValThrGluArgThrAlaGly  139

Qy  2186  ------------------------------CTTCCTGGTGTCCAGTCCGGATCCCACTTCAGATCTCA  2224
                                          |||  |||||:::
Db                                                                          
```

; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 185
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-185

```
Alignment Scores:
Pred. No.:               17.8
Score:                   52.00
Percent Similarity:      44.44%
Best Local Similarity:   31.48%
Query Match:             1.02%
DB:                      1
```

us-09-989-920-100 (1-2754) x US-09-989-920-185 (1-489)

```
Length:          489
Matches:         17
Conservative:    7
Mismatches:      22
Indels:          9
Gaps:            2
```

```
Qy  2144 GGCCAGGGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCC 2203
                                                           ::
Db    27 GlyAsnValAlaThrAsnThrAspGlyLysAsnTyrCysGlyAsnGluGln          46

Qy  2204 GGATCCCACTTGCAGAT-CTCATGCTCTCAGATAGGTGGACAAGTCTTTTGTCACAGT 2262
              :::::: |||  ||||||
Db    47 AlaCysLysIleLysSerPheTyrLeu-----LysTrpAspPhePhe----------- 60

Qy  2263 GCTGGCTCTGTCCTGAGGCCTCATTGCTGGGTGGGTGTGCTC 2304
Db    61 -----AlaLeuLysAsnIleHisCysTrpLysProValLeu 72
```

RESULT 32
US-09-989-920-212
; Sequence 212, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 212
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-212

```
Alignment Scores:
Pred. No.:               18.6
Score:                   51.50
Percent Similarity:      39.16%
Best Local Similarity:   19.44%
Query Match:             1.04%
DB:                      1
```

us-09-989-920-100 (1-2754) x US-09-989-920-212 (1-610)

```
Length:          610
Matches:         49
Conservative:    27
Mismatches:      83
Indels:          93
Gaps:            11
```

```
Qy  1962 CCTGGAGTGAAGACACCAGGGGTG----------CACAGAGGCTGGACAAGCCACAGG 1915
              ||||||  |||  |||||||
Db   382 ProGly---LysThrProHisThrCysAspGlyLysGlyLysAlaPhePheSerSerArg 400

Qy  1914 GGGGCTGGATGGCCACAGCCGGACCATCTCCGGTGGGACTCCTTCCATTGAG 1855
              ||| |||  :::|||||||
Db   401 ThrLeuIleSerHisLysArgValHis---------LeuGlyGluLysProPheLys 416

Qy  1854 GCTGTCCATCCTCCAATGGCAGATGGCAGAATTCTCAGGCCCAGGATACCCTTATGGAGGG 1795
```

```
Db   140 ThrCysLeuTyrPheSerAlaValProGlyIleLeuProSerSerGlnProProIleSer 159
Qy  2225 TGCTCTCAGATAGGTGGACAAGTTCTTTGTCACAGTGCTGGCTCTGTCCTGAGGCCTC 2284
              ||||  |||  :::   ::::::
Db   160 CysSerGluGluGlyAlaGlyAlaThrLeuSerProArg------------------- 173
Qy  2285 ATTGCTGGCTGGGGTGCTCTGCTGGGAAAAGTTTGGGGGCTTGCT---------TGG 2335
              |||  :::   |||  ||||||
Db   174 ---MetGlyGluGluCysValSerValTrpSerHisGluGlyLeuValLeuThrLysLeu 192

Qy  2336 TTAACCACAGAAGAG 2350
              |||||||||||||
Db   193 LeuThrSerGluGlu 197
```

RESULT 30
US-09-989-920-208
; Sequence 208, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 208
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-208

```
Alignment Scores:
Pred. No.:               17.7
Score:                   52.00
Percent Similarity:      43.18%
Best Local Similarity:   31.82%
Query Match:             1.05%
DB:                      1
```

us-09-989-920-100 (1-2754) x US-09-989-920-208 (1-130)

```
Length:          130
Matches:         14
Conservative:    5
Mismatches:      17
Indels:          8
Gaps:            1
```

```
Qy   725 TTTGGTTTCCCCATCGACCGCCAGGCTGGCCAGCTGGGCCCAAGCACTGACCGTCCTTCCCA 666
              |||||||| |||:::|||
Db    53 PheGlyPheProVal-------------------------------GluLeuProArgProGlyPro 64

Qy   665 CCTTATCCCATCAGCTGCCCTGCCACAGGAGCAGAAAACCCATGCCTCAGAGCTACTTCCT 606
              ::: |||||| :::||||||||
Db    65 ThrGlyAlaTyrLysLysValLysAsnGlnThrThrSerSerGluLeuLeuArg 84

Qy   605 AAGGACACTGCC 594
              ||| |||:::|
Db    85 LysGlnThrSer 88
```

RESULT 31
US-09-989-920-185
; Sequence 185, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500

Db 417 CysValGlu----------------------------CysGlyLysSerPheSerTyr------ 426
            ||| :: ; |||

Qy 1794 CCGGGGGGGGCGGCCCCCTCAAAACCACCAAGCTGAGGCCACTGTGAGCAGTCCCCTTTGTAG 1735
Db 427 -------------------------------------SerSerLeuSer 431
                                                 |||

Qy 1734 CAGGAAGAGCAGCTGGACACATGGAGAGGCTTGCAGGTCCAGAGAAGCCAAAAGGTC 1675
Db 432 GlnHisLysArgIleHisThrGlyLysProTyrTyrValCysAspArgCys--------- 448
            ||| :::|::: |||

Qy 1674 CTCAGTCCCCAAGGTACTAGGTGGAGGTGAGATCAGGTCGCCCCTCCCAGGACAAGAGTG 1615
Db 449 ------GlyLysAlaPheArgAsnSerSerGlyLeuThrValHisLysArgIle 464
                  |||:::

Qy 1614 TGAGCTGGCCAGGAGGCCACTGCTGCCCACACTGCTCAAGTCCTTCCACGGCATCAA 1555
Db 465 HisThrGlyGluLysProTyrGluCysAspGluCysGlyLysAlaTyrIleSerHis--- 483
            ||| :::::| :::|||||:::

Qy 1554 CCCTGAGGGACCTGGCCAGGGGATGCAGACCAAGGCCCGGGTCCCCAGGAACAGGA 1495
Db 484 -----------------------------SerSerLeuIleAsnHis 489
                                         |||

Qy 1494 AGCAGGCGGCAAGAAGGAAAGCAA------------------------------ 1471
            :::||||||||

Db 490 LysSerValHisGlnGlyLysGlnProTyrAsnCysGluCysGlyLysSerPheAsnTyr 509

Qy 1470 AGGGCTCTGCATTCCCACCCACGGCGGGGGGTGGACTTCTGCAGGAGCCAGCTGAAGCACCA 1411
Db 510 ArgSerValLeuAspGlnHisLysArgIleHisThrGlyLysLysPro--------- 525
            :::|::: |||

Qy 1410 GAACCTTCCAAGGGGGGCTCGCCAGCCACACAGGACACCAGATGCGTGATAGGCAGGGTT 1351
Db 526 ------------------------------TyrArgCys------------ 528
                                          |||

Qy 1350 AATGCAGAAAACCGCTCATCGAATGCTTCCCCTGCTTCCAAGCGCGAAACCAAGTCATTTT 1291
Db 529 ---AsnGluCysAlaHisIleProAsnAlaThrAlaAspLeuMetLysValAspHis--- 546
               |||

Qy 1290 ATTCTTTAAAAAAGCCCTTACTGTCTGAGGCTTTC 1255
Db 547 ------GluGluGlnLeuSerGluProTyr 556
              :::|::: |||

RESULT 33
US-09-989-920-230
; Sequence 230, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 230
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-230

Alignment Scores:
Pred. No.:   Length: 53
Score: 20.9   Matches: 15
Percent Similarity: 50.50%   Conservative: 6
Best Local Similarity: 53.85%   Mismatches: 11
Query Match: 38.46%   Indels: 7
DB: 1.02%   Gaps: 3
                        1

us-09-989-920-100 (1-2754) x US-09-989-920-230 (1-53)

Qy 721 GTTTCCCATCGCCAGGCTGGCCACGTGGGCCTGGGCGCCAAGCACTGACCGGTCCTTCCCACCTT 662
            ::; ||| ||::||| ||||| ||||||||
Db   14 LeuGlnProAlaProSerTrpLeu----LysGlnAlaLeu------HisLeu 27

Qy 661 ATCCCA--TCAGCTGCCCTGCACAGGAGCAGCAGAAAACCATGCCTCAGAGCTACTTC 608
            ||| |||:| :::||| :::|:||
Db   28 SerProLeuSerSerAlaHisTyrArgHisThrProProHisProAlaAsnPhePhe 46

RESULT 34
US-09-989-920-184
; Sequence 184, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 184
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-184

Alignment Scores:
Pred. No.:   Length: 101
Score: 27.6   Matches: 25
Percent Similarity: 49.00   Conservative: 7
Best Local Similarity: 40.51%   Mismatches: 29
Query Match: 31.65%   Indels: 19
DB: 0.99%   Gaps: 5
                        1

us-09-989-920-100 (1-2754) x US-09-989-920-184 (1-101)

Qy 2673 CAGACTTCTCGTCCG--------CACTATGGTACAGTAACAAGTGCAAAAATATGCTTTA 2623
            ||| |||::|::| ||| :: |||||||
Db   7 GlnAlaProArgProPheLeuTyrHisGlyCysTrpValThrSerGlyValThrSerGlySerHisHisLeu 26

Qy 2622 TTT-----------------CAGGTACAAAAACATGGCAGTAG-GACAACTTTGAG 2585
            ||| |||:::: |||
Db   27 PheProSerLeuPheProIleSerGlnMetTrpGlyHisGlyLeuAspAspGlyLeuHis 46

Qy 2584 CTCAAGCCCACCCTC----------CCAGGACCTCTCACCACCCACCTGACCCAT 2540
            ::: ||| :::::||||||||
Db   47 ArgSerPheHisLeuCysGluLysSerGlyGlnSerAlaArgThrHisLeuCys-- 65

Qy 2539 CCAAGGGCCACCACCCCACAGATACTCCACCACCTTAGAAAAGAGTCACCCA 2483
            ||| ::: |||
Db   66 ProGlySerAlaProGlnAsnGln------ProProAlaSerLeuLysGlnLysPro 82

RESULT 35
US-09-989-920-220
; Sequence 220, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21

align100

Mon Nov 29 08:24:57 2004

```
      ; PRIOR APPLICATION NUMBER: 60/252,500
      ; PRIOR FILING DATE: 2000-11-22
      ; NUMBER OF SEQ ID NOS: 284
      ; SOFTWARE: PatentIn version 3.1
      ; SEQ ID NO 128
      ; LENGTH: 128
      ; TYPE: PRT
      ; ORGANISM: Homo sapien
      US-09-989-920-220

      Alignment Scores:
      Pred. No.:
      Score:                   32.4       Length:        128
      Percent Similarity:      36.76%     Matches:        20
      Best Local Similarity:   29.41%     Conservative:    5
      Query Match:             0.97%      Mismatches:     29
      DB:                                 Indels:         14
                                          Gaps:            5

us-09-989-920-100 (1-2754) x US-09-989-920-220 (1-128)

Qy 1980 CAGTGGCTGCGTGTCGCCCTGGAGTGAAGAAGACCACCAGGGGTGCACAGAGGCTGACAAGC 1921
          ||||||  |||                :::||||||
Db   25 GlnTrpPheLeuLeuGlyLeu--LeuLysThrAlaGlyIle---------TrpGluLys    40

Qy 1920 CACAGGGGGCTGGATGGCCACAGCCGGACCTCGGACCCATCTCCCGTGGGACTCCTTCCA 1861
          |||           |||        |||                        :::||||
Db   41 Glu-----------HisHisArgLeuSerGlnHisGlyAsnIleAsnLeuIlePro     55

Qy 1860 TTTGGAGGTGTCCATCCTCCAATGGCAGATGCAGAATTC---------TCAGGCCCA  1813
          ||||||||                              |||         ||||||
Db   56 ---GluLysGlyArgSerProGlnArgTyr:ValArgPheAsnSerPheSerGlyPro   74

Qy 1812 GGATACCCTTATGGAGGGGCCGGGG 1789
          |||          :::  ||
Db   75 GlySerSerPheSerCysSerGly   82


RESULT 36
US-09-989-920-204
; Sequence 204, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; SOFTWARE: PatentIn version 3.1
; NUMBER OF SEQ ID NOS: 284
; SEQ ID NO 204
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-204

Alignment Scores:
Pred. No.:
Score:                   32.1       Length:        252
Percent Similarity:      26.67%     Matches:        36
Best Local Similarity:   20.00%     Conservative:   12
Query Match:             0.95%      Mismatches:     31
DB:                                 Indels:        101
                                    Gaps:           10

us-09-989-920-100 (1-2754) x US-09-989-920-204 (1-252)

Qy  464 AAAGCTCCTCCTGGCTTT---------CCTCCCTCCCCCAATCTATGGGTCACAGCTAAC 514
          ::  ||||||                |||             |||        |||
Db   45 ArgCysProProAspTyrProIleHisProProArgValLysLeuMetThrThrGlyAsn  64
```

```
Qy  515 -------AGATCTGAGGGCAACTGCTGCTGTGCTAGTGGCCAGGCTGCACCTGCCATCCC 565
                 |||                    :::||||:::  |||   :::
Db   65 AsnThrValArgPheAsnProAsnPheTyrArgAsnGlyLysValCysLeu-SerIleLe  84

Qy  566 CGGCTCTGCCACTTTAGGCCTTCTAGAGGCAGTGTCCTTAGGAAGTAGCTCTGAGGCAT 625
          ||                           |||
Db   84 uGlyTyrTrpThr---GlyPro                                        90

Qy  626 GGGTTTTCTGCTCCTCGTGCAGGGCAGCTGATGGGATAAGGTGGGAGGACCGTCAGTGC 685
          |||                                                   ||
Db   91 ----------------------------------Al                         91

Qy  686 TTGGGCCCCAGCT----------------------------------- 698
          ||||||:::||||
Db   91 aTrpSerProAlaGlnSerIleSerValLeuIleSerIleGlnSerLeuMetThrGl    111

Qy  699 -------------GGCCAGCCTGC--GATGGGGAAACCAAACCATGTCCCCAGCGAA 742
                       :::||||||     :::  |||
Db  111 uAsnProTyrHisAsnGluProGlyPheGluGlnGlnGluArgHisPro-----------  126

Qy  743 GGGCCAGAGTGGGAACCTGTCTCCTCCATGCCCTTCGTCCTGAGGAGCCCTGAGGTGGGCAGC 802
          ||   |||   |||                                 :::
Db  127 -GlyAspSerLysLysAsnTyrAsnGluCys--                            135

Qy  803 AGGGGCCAGGGAAGTTTTCAGGCTTCATCAAAGAGAACAACATCCTCAGCTCCGCACC 862
          |||                        |||
Db  136 ----------------------------IleArgHisGluThrIle----------   141

Qy  863 CCTCATCCTGTATCAGCACTTACCGGTGTGTGAC-------------TGCCCTTGT  905
          ||                    :::                     |||||||
Db  142 -----------------------ArgValAlaValCysAspMetGluGlyLysCysProCys  155


RESULT 37
US-09-989-920-259
; Sequence 259, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 259
; LENGTH: 1533
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-259

Alignment Scores:
Pred. No.:                          Length:       1533
Score:                   17         Matches:        28
Percent Similarity:      48.00%     Conservative:   23
Best Local Similarity:   36.43%     Mismatches:     63
Query Match:             20.00%     Indels:         26
DB:                      0.97%      Gaps:            7

us-09-989-920-100 (1-2754) x US-09-989-920-259 (1-1533)

Qy 1778 CAAAACCACCAGCTGAGGCCACTGTGAGCAGTCCCCTTGTAGCAGGAAGAGCAGCTGG 1719
          |||::::||||||     |||                   :::||||
Db  254 GlnSerHisGlnLeuPheAlaIleIleMetGlu-------CysMetLysLysGluLeu 271

Qy 1718 ACACATGGAAGAGAGCTTGCA------GGTTCCAGAGAAGCCAAAAGGTCCTCAGTCCCC 1665
          |||:::||||||     |||         :::  |||          |||
Db  272 ValThrValPheArgMetGlySerGluGlyGlnGlnAspIleMetAlaIleLeuThr   291
```

Qy   1664 AAGGTACTAGGTGAGGTGAGATCAGGCTGCCC----------------CTCCCAGGA 1623
                  :::|||| || :: ||
Db   292 AlaLeuLeuLysGlyThrAsnValSerAlaProAspGlnLeuSerLeuAlaLeuAlaTrp 311

Qy   1622 CAAGAGTGTGAGCTGGCCAGGAGGCCACTGCTGCCACCACTGCTCAAGTCCTTCCA 1563
                  ::::|||||||||||
Db   312 AsnArgValAspIleAlaArgSerGlnIleIlePheValPhe----------GlyPro 326

Qy   1562 CGCATCAACCCTGAGGGA----------CCTGGCCAGGGATGCAGACC---AAAGGC 1518
                  ||| ||| |||
Db   327 HisTrpThrProLeuGlySerLeuAlaProProThrAspSerLysAlaThrGluLysGlu 346

Qy   1517 CCAGGCGGGTCCCCAGGAACAGGAAGCAGGCGCGCAAGAAGAAAGCAAAGGGCTCTGCATT 1458
                  |||||||| ::: |||
Db   347 LysLysProProMetAlaThrThrLysGlyArgGlyArgGlyLysGlyLysGly------- 363

Qy   1457 CCCACCACGAGGCGCGGGTGGACTCTGCAGGAGGCCAGCTGAAGCCAGAACCTTCCAAGG 1398
                  :::::|| || |||||||| ::: |||||
Db   364 -------LysLysLysValLysGlyValGluValGluGluThrAspThrAspProArg 380

RESULT 38
US-09-989-920-234
; Sequence 234, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Yongming
; APPLICANT: Sun, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 234
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-234

Alignment Scores:
Pred. No.:   33.6   Length: 57
Score: 47.50   Matches: 21
Percent Similarity: 32.00%   Conservative: 3
Best Local Similarity: 28.00%   Mismatches: 22
Query Match: 0.94%   Indels: 29
DB: Gaps: 3

us-09-989-920-100 (1-2754) x US-09-989-920-234 (1-57)

Qy   519 CTGAGGGCAACTGCTGCTAGTGGCCAGGCTGCACCTGCCATCCCCGGCTCT------ 572
                  ||| ||||||||| ||| :::
Db   5 LeuSerAlaHisAlaValLeuGlnAlaSerValProLeuAlaValHisValSerProHis 24

Qy   573 GCCACCTTAGGGCCTTCTCGAGAGCCAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTT 632
                  || ||||||||||
Db   25 AlaArgAlaGlyProSer-- 30

Qy   633 CTGCTCCTGTGCAGGGCAGCTGATCGGATAAGGTGGGAAGGAAGACGGTCAGTGTTGGGCC 692
                  ||| ||||
Db   31 ---TrpSerAlaLeuSerLysTrpVal 39

Qy   693 CCAGCTGGCCAGCCTGGCGATGGGGAAACCAAACCATGTCCCCA 737
                  ||| :::::||||
Db   40 TyrAla------GluAlaAspPheGlnSerValSerCysProPro 52

RESULT 39
US-09-989-920-196
; Sequence 196, Application US/09989920
; GENERAL INFORMATION:

APPLICANT: Macina, Roberto
APPLICANT: Recipon, Herve
APPLICANT: Chen, Sei-Yu
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
FILE REFERENCE: DEX-0291
CURRENT APPLICATION NUMBER: US/09/989,920
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/252,500
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 196
  LENGTH: 88
  TYPE: PRT
  ORGANISM: Homo sapien
US-09-989-920-196

Alignment Scores:
Pred. No.:   34.5   Length: 88
Score: 47.50   Matches: 16
Percent Similarity: 33.75%   Conservative: 11
Best Local Similarity: 20.00%   Mismatches: 24
Query Match: 0.94%   Indels: 29
DB: Gaps: 2

us-09-989-920-100 (1-2754) x US-09-989-920-196 (1-88)

Qy   1986 CCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGCTAACAGGCTAGGCTTTGTGTT 2045
                  ||||||| ::::::::
Db   15 ProGlyPheLeuThrHisCysLeuLysSerArgTrp------------------------ 26

Qy   2046 CTGCGCATTGGTCAGCTTCCACTCGATCCTCCCTAAAGCAATGGGAGGCCCCACTA 2105
                  ::::: ||| ||:||
Db   27 --GlnValProSerLeuAsnHisSerCysAlaProSerGly------------------ 41

Qy   2106 GCCCAGTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGTCCCACAGCTACTG 2165
                  |||:::||
Db   42 -------------------------------------ProLysLeuPro 45

Qy   2166 ATGGCCCGACCAGGTTGAGCTTCCTGGTGTTCGCCGGATCCCACTTGCAGATCTCAT 2225
                  :::::|||||| |||
Db   46 SerSerAlaCysHisSerLeuLeuIleIleSerSerAspGlnValCysValMetHis 65

RESULT 40
US-09-989-920-259
; Sequence 259, Application US/09989920
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Yongming
; APPLICANT: Sun, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 259
; LENGTH: 1533
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-259

Alignment Scores:
Pred. No.:   17.3   Length: 1533
Score: 47.50   Matches: 21
Percent Similarity: 37.11%   Conservative: 15
Best Local Similarity: 21.65%   Mismatches: 33

```
Query Match:      0.94%         Indels:      28
DB:               1             Gaps:        4

us-09-989-920-100 (1-2754) x US-09-989-920-259  (1-1533)

Qy   1423  TGGCTCCTGCGAGAGTCCAGCTCACCCCGGCCTCGTGGGAATGCAGAGCCCTTGCTTTCCTTC  1482
           |||   ||||:::||||:::|||      |||   |||::||
Db     22  TrpGlnLeuGluLeu---ProLysLeuIleSerValHisGlyGlyLeuGlnAsnPhe       40

Qy   1483  --------------------------------TTGCCGCCTGCTTCCTGTTCC  1503
                                           |||:::   :::
Db     41  GluMetGlnLnProLysLeuLysGlnValPheGlyLysLeuIleLysAlaAlaMetThr     60

Qy   1504  TGGGAACCCGCTGGGCCTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATGCGT  1563
                                           |||
Db     61  ThrGly--------------------AlaTrpIlePheThrGlyGlyValSerThr       72

Qy   1564  GGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCCTCCTGGCCAGCTCACACTCTTGT  1623
                                |||:::             ||||||
Db     73  Gly----------ValIleSerHisValGlyAspAlaLeuLysAspHisSerSer        87

Qy   1624  CCTGGGAGGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGACTG  1672
                    |||||   :::  :::  |||||||||
Db     88  LysSerArgGlyArgValCysAla-IleGlyIleIleAlaProTrpGlyIle      103
```

RESULT 41

```
Alignment Scores:
Pred. No.:          37.6          Length:          149
Score:              47.00         Matches:         24
Percent Similarity: 32.95%        Conservative:     5
Best Local Similarity: 27.27%     Mismatches:      23
Query Match:        0.93%         Indels:          36
                    1             Gaps:             6

us-09-989-920-100 (1-2754) x US-09-989-920-279 (1-149)

Qy   1336  AGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTCGTGTCCGTGTGGCTGGCGAGCC  1395
           |||:::
Db     26  SerArgValCysLeuArgLeuValLeuSerTrpSerArgValValCysPheTrp------    43

Qy   1396  CCCCTTGGAAGTTCTGTGCTTCAGCTGGCTTCAGCTGAGTCCAGAGTCCACCCCCTGGTG  1455
                                                              :::||
Db     44  ----------------------TrpSerPhe---TrpLeuPheValSer------ValValCys    54

Qy   1456  GGAATGCAGAGCCCTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT------------  1500
           :::   |||                          |||
Db     55  PheValPheSerCysPheValSerLeuLeu------CysCysCysGlyValValArgLeuTyr    72

Qy   1501  ------TCCTGGGACCC-----------------------------  1512
                 |||
Db     73  PheValValSerTrpGlyValPhePheCysAspLeuLeuArgCysTyrAspAsnVal    92
```

```
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 233
;   LENGTH: 107
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-233

Alignment Scores:
Pred. No.:                43.5      Length:         107
Score:                    46.00     Matches:        15
Percent Similarity:       48.57%    Conservative:   2
Best Local Similarity:    42.86%    Mismatches:     16
Query Match:              0.93%     Indels:         2
DB:                       1         Gaps:           2

us-09-989-920-100 (1-2754) x US-09-989-920-233 (1-107)

Qy  220 TGTGGGAGACCCGCTCTAGAG--ACCTTGCCCTGCTCTGGGGACATC---GCCCAGTGC 167
            |||||            |||||||||||||            |||
Db    9 CysAspSerLeuAlaLeuSerProArgLeuGlnCysSerGlyThrIleSerAlaHisCys 28

Qy  166 AATTTACACTGTCTAGAAAAGTAAGGCCACCAGCGTCATCCACT 122
        ||||||            |||:::::||
Db   29 AsnLeuCysLeuLeuGlySerSerAsnProValSerAlaSer 43
```

RESULT 44
US-09-989-920-245
; Sequence 245, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 245
;   LENGTH: 183
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-245

```
Alignment Scores:
Pred. No.:                43        Length:         183
Score:                    46.00     Matches:        18
Percent Similarity:       46.81%    Conservative:   4
Best Local Similarity:    38.30%    Mismatches:     21
Query Match:              0.93%     Indels:         5
DB:                       1         Gaps:           2

us-09-989-920-100 (1-2754) x US-09-989-920-245 (1-183)

Qy  752 ACTCTGGCCCTCGCTGGGGGACATGGTTTCCCCATCGCCAGGCTGGCC----- 699
            ::||||||      ||||||||||
Db   49 SerLeuAlaPheLeuValGluThrGlyPheSerProThrValAlaArgLeuValSerAsn 68

Qy  698 AGCTGGGGCCCAAGCACTGA-CCGTCCTTCCCACCTTATCCATCAGCTGCCCTGCACA 640
        ||||||            |||:::||
Db   69 SerTrpProProValValArgProProLeuProSerGlnSerAlaGlyIle------Thr 86

Qy  639 GGAGCAGAAAACCCATGCCTC 619
        |||      ||||||||
Db   87 GlyValGlyProProCysLeu 93
```

RESULT 45
US-09-989-920-193
; Sequence 193, Application US/09989920
; GENERAL INFORMATION:
;   APPLICANT: Macina, Roberto
;   APPLICANT: Recipon, Herve
;   APPLICANT: Chen, Sei-Yu
;   APPLICANT: Sun, Yongming
;   APPLICANT: Liu, Chenghua
;   TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;   FILE REFERENCE: DEX-0291
;   CURRENT APPLICATION NUMBER: US/09/989,920
;   CURRENT FILING DATE: 2001-11-21
;   PRIOR APPLICATION NUMBER: 60/252,500
;   PRIOR FILING DATE: 2000-11-22
;   NUMBER OF SEQ ID NOS: 284
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 193
;   LENGTH: 63
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-920-193

```
Alignment Scores:
Pred. No.:                46        Length:         63
Score:                    45.50     Matches:        16
Percent Similarity:       39.34%    Conservative:   8
Best Local Similarity:    26.23%    Mismatches:     18
Query Match:              0.90%     Indels:         19
DB:                       1         Gaps:           3

us-09-989-920-100 (1-2754) x US-09-989-920-193 (1-63)

Qy 1820 GAGAATTCTGCATCTCGCCATTGGAGGATGGACAGCCTCAAATGGAGGAGTCCCACGGGA 1879
            ||:::::||||||            :::||
Db   17 GluSerAsnGlnProArgPheGlyGlyThrpGlyThrpGlyAspGlyAlaThr-------- 33

Qy 1880 GATGGGGTCCGAGGTCCGGCTGGGCCATCCAGCGCCCCCGTGGCCTTGTCCAGCCTCTGTGC 1939
        ||:::||            :::||
Db   34 -------------------------PheProProTyrLeuLeuPhePheTyrIleProIleCys 46

Qy 1940 ACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTGC--AGTTCCTTCTTCGT 1996
        |||      :::||      ||||||||
Db   47 ThrLeuArgIle-------------------HisLeuArgSerPheLysArg 58

Qy 1997 GAG 1999
        |||
Db   59 Glu 59
```

Search completed: November 23, 2004, 08:51:46
Job time : 29.5 secs

Exhibit B (handwritten)

```
Qy  2401 GCACGGTTACTGTGTTCTCAAGTTCATGTATTTAAAATGATTTCTTCTAAAGATGTA 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 GCACGGTTACTGTGTTCTCAAGTTCATGTATTTAAAATGATTTCTTCTAAAGATGTA 2460

Qy  2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTCTCAAAGGTGGTGGGGAGTAGTGTC 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTCTCAAAGGTGGTGGAGTATCTGTC 2520

Qy  2521 GGGGTGGTGGTGGCCCTTGGATGGGTCAGGTGGTGGGTGGTGAGAGCCTGGGGAGGTGGGCGT 2580
         |||||||||||||||||||||||||||||||||||
Db  2521 GGGGTGGTGGTGGCCCTTGGATGGGTCAGGTGGTGGGTGGGAGAGGTCCTGGGGAGGTGGGCGT 2580

Qy  2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTGTTGTTACCTGAAATAAAGCATATTTTGCACT 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTGTTGTTACCTGAAATAAAGCATATTTTGCACT 2640

Qy  2641 TGTTACTGTACCATGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGTTGGGTTAGAA 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 TGTTACTGTACCATGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGTTGGGTTAGAA 2700

Qy  2701 GCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 GCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
```

```
RESULT 2
AC079988/c
LOCUS       AC079988      211305 bp    DNA     linear   PRI 09-MAR-2002
DEFINITION  Homo sapiens BAC clone RP11-795C1 from 2, complete sequence.
ACCESSION   AC079988
VERSION     AC079988.6  GI:18873965
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 211305)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 211305)
  AUTHORS   Waligorski,J., Haakenson,W. and Boyer,E.
  TITLE     The sequence of Homo sapiens BAC clone RP11-795C1
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 211305)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-SEP-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 211305)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 211305)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Feb 23, 2002 this sequence version replaced gi:18042517.
            -----------  Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            -----------  Summary Statistics
            Center project name: H_NH0795C01
```

```
NOTICE:   This sequence may not represent the entire insert of this
clone.  It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson,Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR:  pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-60M20, 2000 bp overlap.
Actual start of this clone is at base position 1 of RP11-795C1;
actual end is at base position 111002 of RP11-6906.

Single plasmid region exists between 76851 and 76869.  Polymorphisms
exists between AC067960, AC013399 and AC079988.  Data from AC067960
and AC013399 was used to finish AC079988.

FEATURES             Location/Qualifiers
     source          1..211305
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-795C1"
                     /clone_lib="RPCI-11"
     repeat_region   1..115
     repeat_region   1042..1153
                     /rpt_family="MIR"
     repeat_region   1682..1832
                     /rpt_family="MER1_type"
     repeat_region   1974..2095
                     /rpt_family="(TG)n"
     repeat_region   2096..2144
                     /rpt_family="Alu"
     repeat_region   2145..2301
                     /rpt_family="(TAAA)n"
     repeat_region   2371..2472
                     /rpt_family="Alu"
     repeat_region   2491..2583
                     /rpt_family="MIR"
     repeat_region   3648..3699
                     /rpt_family="MIR"
     repeat_region   3726..3838
                     /rpt_family="T-rich"
     repeat_region   4074..4169
                     /rpt_family="MIR"
     repeat_region   4375..4466
                     /rpt_family="MIR"
     repeat_region   4627..4725
                     /rpt_family="MER1_type?"
     repeat_region   4825..4920
                     /rpt_family="L2"
```

Accession (ai 10/08/10 2)

```
repeat_region    4921..5082
                 /rpt_family="L2"
repeat_region    5111..5368
                 /rpt_family="(TG)n"
repeat_region    5429..5655
                 /rpt_family="(TG)n"
repeat_region    6114..6269
                 /rpt_family="(TG)n"
repeat_region    6325..6431
                 /rpt_family="MIR"
repeat_region    6541..6674
                 /rpt_family="MIR"
                 /rpt_family="MIR"
repeat_region    6713..6936
                 /rpt_family="MaLR"
repeat_region    6937..7459
                 /rpt_family="ERV1"
repeat_region    7460..7663
                 /rpt_family="MaLR"
repeat_region    7679..7869
                 /rpt_family="MIR"
repeat_region    7892..8092
                 /rpt_family="MaLR"
repeat_region    8693..8775
                 /rpt_family="MIR"
repeat_region    8821..8897
                 /rpt_family="MIR"
repeat_region    9061..9445
                 /rpt_family="L2"
repeat_region    9608..9721
                 /rpt_family="(TA)n"
repeat_region    9723..9772
                 /rpt_family="(GA)n"
repeat_region    9904..10068
                 /rpt_family="MER1_type"
repeat_region    10147..10247
                 /rpt_family="L2"
repeat_region    10496..10921
                 /rpt_family="MaLR"
repeat_region    11078..11199
                 /rpt_family="MIR"
repeat_region    11298..11841
                 /rpt_family="MaLR"
repeat_region    12041..12220
                 /rpt_family="MIR"
repeat_region    12379..12465
                 /rpt_family="MIR"
repeat_region    12543..13210
                 /rpt_family="L1"
repeat_region    14121..14169
                 /rpt_family="L2"
repeat_region    14163..14729
                 /rpt_family="L2"
repeat_region    14802..14966
                 /rpt_family="MER1_type"
repeat_region    14974..15129
                 /rpt_family="L2"
repeat_region    15286..15477
                 /rpt_family="ERVL"
repeat_region    15518..15888
                 /rpt_family="L1"
repeat_region    15889..16196
                 /rpt_family="Alu"
repeat_region    16197..16848
                 /rpt_family="L1"
repeat_region    16861..16884
                 /rpt_family="AT_rich"
repeat_region    17160..17297
                 /rpt_family="L1"
repeat_region    17298..17591
                 /rpt_family="Alu"
repeat_region    17592..18212
                 /rpt_family="L1"
```

```
repeat_region    18213..18522
                 /rpt_family="Alu"
repeat_region    18523..19061
                 /rpt_family="L1"
repeat_region    19062..19198
                 /rpt_family="L1"
repeat_region    19199..19512
                 /rpt_family="ERVL"
                 /rpt_family="Alu"

Query Match           98.7%;  Score 2718.4;  DB 9;  Length 211305;
Best Local Similarity 99.8%;  Pred. No. 0;
Matches 2722;  Conservative 6;  Indels 0;  Gaps 0;

Qy      1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGGAGATGACTGCTGTCCCTTCGCATTTGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  90196 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGGAGATGACTGCTGTCCCTTCGCATTTGG 90137

Qy     61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGTGTGACCTCCTGGA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  90136 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGTGTGACCTCCTGGA 90077

Qy    121 GAGTGGATGACGCTGGTGGCCTTAGCTTTCTAGACAGTGTAAATTGACTGGGCGATGT 180
          ||||||||||    |||||||||||||||||||||||||||   |||  |||||  ||
Db  90076 GAGTGGATGAGAGCTGTGGCCTTAGCTTTTCTAGACAGTGTAAGTGAAGTGGGCAAGGC 90017

Qy    181 CCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  90016 CCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 89957

Qy    241 CTTCCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89956 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 89897

Qy    301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCATCGCCCTCCTGGGACCTGCCCAGGGCACCCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89896 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCATCGCCCTCCTGGGACCTGCCCAGGGCACCCT 89837

Qy    361 GGCCAGTGAGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGGCACCCT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89836 GGCCAGTGAGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGGCACCCT 89777

Qy    421 GAGAGCTCTGAAACCCCACTTAGCTTCCAGACCTTTCTGCAAAAGTCCTCCTGGCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89776 GAGAGCTCTGAAACCCCACTTAGCTTCCAGACCTTTCTGCAAAAGTCCTCCTGGCTT 89717

Qy    481 TCCTCCTCCCCAATCTATGGGTCACAGTCAACAGATCTGAGGCAACTGCTGTGCTAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89716 TCCTCCTCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGCAACTGCTGTGCTAG 89657

Qy    541 TGGCCAGGGTGCACCTGCCATCCCGGCTGCCACTTAGGGCCTTCAGAGAGGCAGTG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89656 TGGCCAGGGTGCACCTGCCATCCCGGCTGCCACTTAGGGCCTTCAGAGAGGCAGTG 89597

Qy    601 TCCTTAGGAGTAGCTCTGAGGCATGGGTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89596 TCCTTAGGAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 89537

Qy    661 TAAGGTGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGCCAGCTGGGGATGGGGAAA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89536 TAAGGTGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGCCAGCTGGGGATGGGGAAA 89477

Qy    721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89476 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 89417

Qy    781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCAGCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89416 GAGGAGCCCTGAGGTGGGCAGCAGGGGCAGCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 89357

Qy    841 ACAACATCCCTCAGCTCCGCACCCCTCATCCTGTATCCGGTGTGTGACTGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89356 ACAACATCCCTCAGCTCCGCACCCCTCATCCTGTATCCGGTGTGTGACTGCC 89297
```

```
Qy    901 CTTGTCAGCTAGCATACGGTGGCCCACCTGGCTGTGTTATGCCACTGATTTA  960
          ||||||||||||||||||||||||||||||||||||| || |||||||||||
Db  89296 CTTGTCAGCTAGCATACGGTGGCCCACCTGGCTGTGTTATGCCACTGATTTA  89237

Qy    961 TGATAGGGAAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCATCACAAAAAAA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89236 TGATAGGGAAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCATCACAAAAAAAA  89177

Qy   1021 AATTCTATTTTTAGTAGACATGTATTTACCAAAAAATATGTACTCAATTATTGTATTTTG  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89176 AATTCTATTTTTAGTAGACATGTATTTACCAAAAAATATGTACTCAATTATTGTATTTTG  89117

Qy   1081 GATTTTATCAATTTAAAAATTGTGGAAAATTGTTTGCTCTTACGCCAACATAATATTGAT  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89116 GATTTTATCAATTTAAAAATTGTGGAAAATTGTTTGCTCTTACGCCAACATAATATTGAT  89057

Qy   1141 TTTGCCTCTTGGCTCTGAAAGCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAAGTC  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89056 TTTGCCTCTTGGCTCTGAAAGCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAAGTC  88997

Qy   1201 TGCTGACTACTGAGCCAGACCTTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88996 TGCTGACTACTGAGCCAGACCTTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC  88937

Qy   1261 CTCAGACAGTAAGGGCTTTTTTAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88936 CTCAGACAGTAAGGGCTTTTTTAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG  88877

Qy   1321 GGAAGCATTCAGATGAGCGGTTTCTGATTAACCCGCCATCACGCATCCGTGTCCTG  1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88876 GGAAGCATTCAGATGAGCGGTTTCTGATTAACCCGCCATCACGCATCCGTGTCCTG  88817

Qy   1381 TGTGGCTGGCGAGCCCCCCTTGGAAGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88816 TGTGGCTGGCGAGCCCCCCTTGGAAGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA  88757

Qy   1441 CCCCGCCCTCGTGTGGGAATGCAGAGCCCTTTGCTTTCCTTTGCGCCCTGCTTCCTGT  1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88756 CCCCGCCCTCGTGTGGGAATGCAGAGCCCTTTGCTTTCCTTTGCGCCCTGCTTCCTGT  88697

Qy   1501 TCCTGGGACCCGCTGGGCCTTTGGTCTGCATCCCGTGGCCAGGTCCCTCAGGGGTTGATG  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88696 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCGTGGCCAGGTCCCTCAGGGGTTGATG  88637

Qy   1561 CGTGGGAGGAGGACTTTGAGCAGTGGGTGGGCAGCAGTGGCCTTCCTGCCAGCTCACACTCT  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88636 CGTGGGAGGAGGGACTTTGAGCAGTGGGTGGGCAGCAGTGGCCTTCCTGCCAGCTCACACTCT  88577

Qy   1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGACTGAGGACCTT  1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88576 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGACTGAGGACCTT  88517

Qy   1681 TTGGCTTCTCTGGAGCCTGCAGCCTCTTCCCAGCTGCTCCAGCTCGTCTCTCTGCTACAA  1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88516 TTGGCTTCTCTGGAGCCTGCAGCCTCTTCCCAGCTGCTCCAGCTCGTCTCTCTGCTACAA  88457

Qy   1741 AGGGGACTGCTCACAGTGCCCTCAGCTTGGTGTTTTGAGGGGCCGCCCCCGGCCCCTCC  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88456 AGGGGACTGCTCACAGTGCCCTCAGCTTGGTGTTTTGAGGGGCCGCCCCCGGCCCCTCC  88397

Qy   1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCATTGGAGGATGGACAGCCTCAAA  1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88396 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCATTGGAGGATGGACAGCCTCAAA  88337

Qy   1861 TGGAAGGAGTCCCACGGAGGGAGATGGGTCCGAGTCCGGCTGTGGCCATCCAGCCCCTGTG  1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88336 TGGAAGGAGTCCCACGGAGGGAGATGGGTCCGAGTCCGGCTGTGGCCATCCAGCCCCTGTG  88277

Qy   1921 GCTTGTCCAGCCTCTGTCACCCCTGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG  1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88276 GCTTGTCCAGCCTCTGTCGCCCCTGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG  88217

Qy   1981 CAGTTCCTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGCTAACAGGCTAGGCTTT  2040
```

```
Db  88216 CAGTTCCTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGCTAACAGGCTAGGCTTT  88157
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2041 GTGTTCTGCGATTGGTCAGCTTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCCC  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88156 GTGTTCTGCGATTGGTCAGCTTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCCC  88097

Qy   2101 CACTAGCCCAGTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88096 CACTAGCCCAGTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC  88037

Qy   2161 TACTGATGGCCGAGCCAGGCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT  2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  88036 TACTGATGGCCGAGCCAGGCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT  87977

Qy   2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCTCCTGAGG  2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  87976 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCTCCTGAGG  87917

Qy   2281 CCTCATTGCTGGCTGGGTGCTCTGCTGGGAGGTTGCGGGGCTTGGTTGGTTAAC  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  87916 CCTCATTGCTGGCTGGGTGTCTGCTGGGAGGTTGCGGGGCTTGGTTGGTTAAC  87857

Qy   2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCCTGGCAGCCTCCCCGTGCTGGGTGGA  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  87856 CACAGAGAGAGGGGACTGTTGGGTGCTCCTCTCCCGTGCTGGGTGGA  87797

Qy   2401 GCACGGTTACTGTGTTCCTAATGTCATGTAATTAAAATGATTTCTTTCTAAGATGTA  2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  87796 GCACGGTTACTGTGTTCCTAATGTTCATGTAATTAAAATGATTTCTTTCTAAGATGTA  87737

Qy   2461 ACCTCCACACCTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  87736 ACCTCCACACCTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC  87677

Qy   2521 GGGGTGCGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCTCGGCGAGGTGGGCGT  2580
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  87676 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTCAGGTGGGTGGGAGGTCCTGGGGAGGTGGGCGGT  87617

Qy   2581 TGAGCTCAAAGTGTCTACTGCCATGTTTTGTACCTGAAATAAAGCATATTTGCACT  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  87616 TGAGCTCAAAGTGTCTCCTACTGCCATGTTTTGTACCTGAAATAAAGCATATTTGCACT  87557

Qy   2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGCTTGTTAGAA  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  87556 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGCTTGTTAGAA  87497

Qy   2701 TGCAAATAAAACTCACATTTGTAAGAAA  2728
          ||||||||||||||||||||||||||||
Db  87496 TGCAAATAAAACTCACATTTGTAAGAAA  87469
```

RESULT 3
AC140107/c

```
LOCUS       AC140107              225203 bp    DNA     linear   HTG 21-FEB-2003
DEFINITION  Homo sapiens chromosome 17 clone RP11-1201C10 map 17, WORKING DRAFT
            SEQUENCE, 4 unordered pieces.
ACCESSION   AC140107
VERSION     AC140107.1  GI:28460978
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 225203)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-1201C10
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 225203)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chrobel,Y.,,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
```

Exhibit B (page 1 of 2) (handwritten)

```
            GenCore version 5.1.6
     Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 29, 2004, 08:42:07 ; Search time 227 Seconds
                                     (without alignments)
                          8623.403 Million cell updates/sec

Title:         US-09-989-920-100
Perfect score: 2754
Sequence:      1 gccagagcagcctcagctt........aaaaataaaagatgcggcc 2754

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:  824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters:    1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued_Patents NA:*
      1: /cgn2_6/ptcdata/1/ina/5A_COMB.seq:*
      2: /cgn2_6/ptcdata/1/ina/5B_COMB.seq:*
      3: /cgn2_6/ptcdata/1/ina/6A_COMB.seq:*
      4: /cgn2_6/ptcdata/1/ina/6B_COMB.seq:*
      5: /cgn2_6/ptcdata/1/ina/PCTUS_COMB.seq:*
      6: /cgn2_6/ptcdata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1747.6	63.5	1853	3	US-09-232-160-3	Sequence 3, Appli
	2	52	1.9	7218	3	US-08-233-463-14	Sequence 14, Appl
	3	49	1.8	6330	3	US-09-880-427-2	Sequence 2, Appli
	4	49	1.8	6330	3	US-09-306-538B-2	Sequence 2, Appli
	5	49	1.8	19806	4	US-09-740-028A-3	Sequence 3, Appli
	6	49	1.8	19806	4	US-10-118-037-3	Sequence 3, Appli
	7	47	1.7	265	4	US-09-513-999C-27278	Sequence 27278, A
c	8	42.6	1.5	640681	4	US-09-790-988-1	Sequence 1, Appli
	9	41.6	1.5	210	4	US-09-248-796A-12137	Sequence 12137, A
c	10	41.4	1.5	329	4	US-09-513-999C-10733	Sequence 10733, A
c	11	40.8	1.5	318	4	US-09-513-999C-11322	Sequence 11322, A
	12	40.8	1.5	398	4	US-09-513-999C-23077	Sequence 23077, A
	13	40.4	1.5	1141	4	US-09-806-708B-22	Sequence 22, Appli
	14	40.4	1.5	640681	4	US-09-790-988-1	Sequence 1, Appli
	15	39.6	1.4	1107	4	US-09-248-796A-8906	Sequence 8906, Ap
c	16	39.4	1.4	588	1	US-09-328-111-321	Sequence 321, Appl
c	17	39.4	1.4	5852	1	US-07-867-106-2	Sequence 2, Appli
	18	39	1.4	303	4	US-09-601-198-179	Sequence 179, Appl
	19	38.8	1.4	9636	1	US-08-323-170B-1	Sequence 1, Appli
	20	38.8	1.4	9636	3	US-08-954-441-1	Sequence 1, Appli
c	21	38.4	1.4	2422	1	US-07-867-106-5	Sequence 5, Appli
c	22	38.4	1.4	5852	2	US-07-867-106-2	Sequence 2, Appli
c	23	38.4	1.4	54945	4	US-09-967-669-10	Sequence 10, Appl
c	24	38.2	1.4	1099	4	US-09-270-767-11371	Sequence 11371, A
c	25	38.2	1.4	3009	4	US-09-248-796A-6806	Sequence 6806, Ap
c	26	38.2	1.4	3454	4	US-09-270-767-11909	Sequence 11909, A
c	27	38.2	1.4	5185	3	US-08-971-395-5	Sequence 5, Appli
c	28	38.2	1.4	8561	3	US-09-112-450-3	Sequence 3, Appli
c	29	38.2	1.4	8561	4	US-09-419-291A-3	Sequence 3, Appli
c	30	38.2	1.4	98844	4	US-09-791-211-10	Sequence 10, Appl
c	31	38	1.4	2007	2	US-08-743-637B-169	Sequence 169, App
c	32	38	1.4	2007	3	US-08-526-840B-169	Sequence 169, App
c	33	38	1.4	2028	3	US-09-134-001C-1710	Sequence 1710, Ap
c	34	37.8	1.4	4140	3	US-08-894-731-2	Sequence 2, Appli
c	35	37.8	1.4	8920	3	US-08-446-855A-1	Sequence 1, Appli
c	36	37.8	1.4	8920	3	US-09-150-741-1	Sequence 1, Appli
c	37	37.8	1.4	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c	38	37.8	1.4	1664976	4	US-09-692-570-1	Sequence 1, Appli
c	39	37.4	1.4	2663	1	US-08-136-743B-3	Sequence 3, Appli
c	40	37.4	1.4	6040	4	US-10-204-708-70	Sequence 70, Appl
c	41	37.4	1.4	74962	4	US-09-685-853A-3	Sequence 3, Appli
c	42	37.2	1.4	942	4	US-09-270-767-3767	Sequence 3767, Ap
c	43	37.2	1.4	942	4	US-09-270-767-19049	Sequence 19049, A
c	44	37.2	1.4	1928	4	US-09-674-82 6B-5	Sequence 5, Appli
c	45	36.8	1.3	80246	3	US-09-078-294-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-09-232-160-3
; Sequence 3, Application US/09232160
; Patent No. 6368794
; GENERAL INFORMATION:
; APPLICANT: Steve Daniel
; APPLICANT: James Gilmore
; APPLICANT: Susan G. Stuart
; APPLICANT: Laura Stuve
; TITLE OF INVENTION: DETECTION OF ALTERED EXPRESSION OF GENES REGULATING CELL
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: PA-0003 US
; CURRENT APPLICATION NUMBER: US/09/232,160
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 1283330
US-09-232-160-3

Query Match        63.5%;  Score 1747.6;  DB 3;  Length 1853;
Best Local Similarity   99.8%;  Pred. No. 0;
Matches 1750;  Conservative  0;  Mismatches  0;  Indels  4;  Gaps  0;

Qy   994 TGTTTTCTCCCCATCACAAAAAAAAAATTCTTATTTTAGTAGACATGTATTTACCAA 1053
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TGTTTTCTCCCCATCACAAAAAAAAAAAATTCTTATTTTAGTAGACATGTATTTACCAA 60

Qy  1054 AAATAGTACTCAATTATTGTATTTTGGATTTATCAATTAAAATTGTGGAAATTTGT 1113
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AAATATGTACTCAATTATTTGGATTTTGGATTTATCAATTAAAATTAAATTGTGGAAATTTGT 120

Qy  1114 TTGCTCTTAGCCAACATATATTGATTTGCCTCTTGGCTCTGACTACTGAGC 1173
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTGCTCTTAGCCAACATATATTGATTTGCCTCTTGGCTCTGACTACTGAGC 180

Qy  1174 ACCGTCTAGCCCGTTACAGAAAAGTCTGCTGCTGAGCTACTGAGCACTACCTCC 1233
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACCGTCTAGCCCGTTACAGAAAAGTCTGCTGCTGAGCTACTGAGCACTACCTCC 240

Qy  1234 ATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTTTAAAGAATAAA 1293
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTTTTAAAGAATAAA 300

Qy  1294 ATGACTTGGTTTGCCGTTGGCCTTGGAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAAC 1353
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   2434 TTAAAATGATTTCTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCT 2493
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TTAAAATGATTTCTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCT 1500

Qy   2494 TTTCTAAAGTGGTGGGAGTATCGTCGGGGTGGGTGTGGGGCCTTGGATGGGTCAGTGGG 2553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TTTCTAAAGTGGTGGGAGTATCGTCGGGGTGGGTGTGGGGCCTTGGATGGGTCAGTGGG 1560

Qy   2554 TGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTGAAGAGTTGTCCTACTGCCATGTTTTG 2613
          ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||| |
Db   1561 TGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTGTCCTACTGCCATGTTTTTG 1620

Qy   2614 TACCTGAAATAAAGCATATTTGCACTTGTTACTGTACCATAGTGCGGACGAAGTCTG 2673
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TACCTGAAATAAAGCATATTTGCACTTGTTACTGTACCATAGTGCGGACGAAGTCTG 1680

Qy   2674 TATGTGGGATTGTGTCTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAGAAAAAAA 2733
          ||||||||| |||| ||||||||||||||||||||||||||||||||||||||
Db   1681 TATGTGGGATCGTGCTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAGACAAAAAA 1740

Qy   2734 AAAAAATAAAAAGA 2747
          ||||| ||| ||||
Db   1741 AAAAACAAAAGAGA 1754
```

RESULT 2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
GENERAL INFORMATION:
  APPLICANT: DORNER, F.
  APPLICANT: SCHEIFLINGER, F.
  APPLICANT: FALKNER, F.G.
  TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
  NUMBER OF SEQUENCES: 52
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Foley & Lardner
    STREET: 800 Diagonal Road, Suite 500
    CITY: Alexandria
    STATE: VA
    COUNTRY: USA
    ZIP: 22313-0299
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/232,463
    FILING DATE:
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US/07/935,312
    FILING DATE:
    APPLICATION NUMBER: EP 91 114 300.6
    FILING DATE: 26-AUG-1991
  ATTORNEY/AGENT INFORMATION:
    NAME: BENT, Stephen A.
    REGISTRATION NUMBER: 29,768
    REFERENCE/DOCKET NUMBER: 30472/114 IMMJ
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (703)836-9300
    TELEFAX: (703)683-4109
    TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 7218 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  IMMEDIATE SOURCE:
    CLONE: pTZgpt-F1s

```
Db    301 ATGACTTGGTTGCGCTTGGAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAAC 360

Qy   1354 CCTGCCTATCACGCATCTCGTGTCCTGTGGCTCGGCGGAGCCCCCTTTGGAAGGTTCTGG 1413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCTGCCTATCACGCATCTCGTGTCCTGTGGCTCGGCGGAGCCCCCTTTGGAAGGTTCTGG 420

Qy   1414 TGCTTCAGTGGCTCCTGCAGAGTCCACCCGCCTCGTGTGGGAATGCAGAGCCCTTTG 1473
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGCTTCAGTGGCTCCTGCAGAGTCCACCCGCCTCGTGTGGGAATGCAGAGCCCTTTG 480

Qy   1474 CTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCCTGGGCCTTGGTCTGCATC 1533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCCTGGGCCTTGGTCTGCATC 540

Qy   1534 CCCTGGCCAGTCCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGC 1593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCCTGGCCAGTCCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGC 600

Qy   1594 AGTGGCCTCCTGGCCAGCTCACACCTCTGTCCTGGAGGGCAGCCTGATCTCACCTCCA 1653
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGTGGCCTCCTGGCCAGCTCACACCTCTGTCCTGGAGGGCAGCCTGATCTCACCTCCA 660

Qy   1654 CCTAGTACCTTGGGGACTGAGGACCTTTGCTTCTCTGGAGCCTGCAAGCCCTCTTCCA 1713
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCTAGTACCTTGGGGACTGAGGACCTTTGCTTCTCTGGAGCCTGCAAGCCCTCTTCCA 720

Qy   1714 TGTGTCCAGTCAGTCGCTCTTCCTGCTCTGCTCACAGTGGCCTCAGTTGGTGG 1773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGTGTCCAGTCAGTCGCTCTTCCTGCTCTGCTCACAGTGGCCTCAGTTGGTGG 780

Qy   1774 TTTTGAGGGGCGGCCGCCCCCCGGCCCTTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATC 1833
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTTTGAGGGGCGGCCGCCCCCCGGCCCTTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATC 840

Qy   1834 TGCCATTGGAGGATGACAGCCTCAAATGAAGGAGTCCACGGGAGATGGGTCCGAGGT 1893
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCATTGGAGGATGACAGCCTCAAATGAAGGAGTCCACGGGAGATGGGTCCGAGGT 900

Qy   1894 CCGGCTGTGGCCATCAGCCCCTGGCTTGTCCAGCGCCTCTGTGCACCCCTGTGTCTT 1953
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCGGCTGTGGCCATCAGCCCCTGGCTTGTCCAGCGCCTCTGTGCACCCCTGTGTCTT 960

Qy   1954 CACTCCAGGGGCAGACAGCAGACAGCCACTGCAGTTCCTTCTTCGTGAGTAACAGTAGTGATA 2013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CACTCCAGGGGCAGACAGCAGACAGCCACTGCAGTTCCTTCTTCGTGAGTAACAGTAGTGATA 1020

Qy   2014 GCAGTGGGGCGGCTAACAGGGCTAGGCTTGTGTTCTGCATTTGGTCAGCTTCTCCACTCGA 2073
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCAGTGGGGCGGCTAACAGGGCTAGGCTTGTGTTCTGCATTTGGTCAGCTTCTCCACTCGA 1080

Qy   2074 TCCTCCCTAAAGCAATGGGGAGGGCCCCCACTAGCCCAGTTTCAGGAAGTCAACTCGGGAG 2133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCCTCCCTAAAGCAATGGGGAGGGCCCCCACTAGCCCAGTTTCAGGAAGTCAACTCGGGAG 1140

Qy   2134 GTTAGATGGGGCCAGGGTCCCACAGCTACTGATGGCCCAGCCAGGTTGAGCTTCCTGG 2193
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GTTAGATGGGGCCAGGGTCCCACAGCTACTGATGGCCCAGCCAGGTTGAGCTTCCTGG 1200

Qy   2194 TGTCCAGTTCCGATCCCACTTGCAGATCTCAGATCTCATGCTCAGATAGTGGGACAAGTTCTTT 2253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TGTCCAGTTCCGATCCCACTTGCAGATCTCAGATCTCATGCTCAGATAGTGGGACAAGTTCTTT 1260

Qy   2254 TGTCACAGTGCTGGTCTGTCTCTGCCTCATTGCTGGCTCGGGTGTCGCTCTGCTGGGGAA 2313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGTCACAGTGCTGGTCTGTCTCTGCCTCATTGCTGGCTCGGGTGTCGCTCTGCTGGGGA 1320

Qy   2314 AAGCTTTGCGGGGCTTGCGGGGCTTGCTTGGTTAACCACAGAAGGGGACTGTTTGGGGTGCCTC 2373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AAGCTTTGCGGGGCTTGCGGGGCTTGCTTGGTTAACCACAGAAGGGGACTGTTTGGGGTGCCTC 1380

Qy   2374 TCTGCAGCCTCCCCCGTGCGGGGTGGGTCGGTGGGTTACTGTGTTTCTCTAATGTTCATGTAT 2433
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TCTGCAGCCTCCCCCGTGCGGGGTGGGTCGGTGGGTTACTGTGTTTCTCTAATGTTCATGTAT 1440
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: December 6, 2004, 14:03:15 ; Search time 627 Seconds
(without alignments)
3.920 Million cell updates/sec

Title: US-09-989-920-100
Perfect score: 2754
Sequence: 1 gccaggagcagcctcagctt........aaaaataaaaagatgcggcc 2754

Scoring table: IDENTITY.NUC
Gapop 10.0 , Gapext 0.5

Searched: 3 seqs, 446265 residues

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : *txt:*
1: /staff_overflow/sdavid-tmp/dec04/yu920/10198585.txt:*
2: /staff_overflow/sdavid-tmp/dec04/yu920/10765092.txt:*
3: /staff_overflow/sdavid-tmp/dec04/yu920/10800346.txt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	2604.1	94.6	208378	3	AC079988	ACCESSION:AC079988
c 2	1821.8	66.2	187722	2	AC079988	ACCESSION:AC079988
3	623.101	22.6	187722	2	AC079988	ACCESSION:AC079988
4	29.6	1.1	208378	3	AC079988	ACCESSION:AC079988
5	24.4	0.9	50165	1	AC079988	ACCESSION:AC079988
c 6	23.8	0.9	50165	1	AC079988	ACCESSION:AC079988

ALIGNMENTS

RESULT 1
AC079988/c
LOCUS   AC079988   208378 bp   DNA   linear   HTG 14-OCT-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-795C1, WORKING DRAFT SEQUENCE,
32 unordered pieces.
ACCESSION   AC079988
VERSION   AC079988.3 GI:10800346
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 208378)
AUTHORS   Waterston,R.H.
TITLE   The sequence of Homo sapiens clone
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 208378)
AUTHORS   Waterston,R.H.
TITLE   Direct Submission

JOURNAL   Submitted (20-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
[WARNING] On Nov 2, 2001 this sequence was replaced by a newer
version gi:16596673.
On Oct 13, 2000 this sequence version replaced gi:10765092.

COMMENT   -------------- Genome Center --------------
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
--------------- Project Information ---------------
Center project name: H_NH079SC01
-------------- Summary Statistics --------------
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 189029 bases at least Q40
Consensus quality: 197040 bases at least Q30
Consensus quality: 200793 bases at least Q20
Insert size: 217000; agarose-fp
Insert size: 205278; sum-of-contigs
Quality coverage: 3.74 in Q20 bases; agarose-fp
Quality coverage: 3.99 in Q20 bases; sum-of-contigs
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*   1   1933: contig of 1933 bp in length
*   1934   2033: gap of unknown length
*   2034   4225: contig of 2192 bp in length
*   4226   4335: gap of unknown length
*   4326   6989: contig of 2664 bp in length
*   6990   7089: gap of unknown length
*   7090   9679: contig of 2590 bp in length
*   9780   9779: gap of unknown length
*   9780   12591: contig of 2812 bp in length
*   12592   12691: gap of unknown length
*   12692   15210: contig of 2519 bp in length
*   15211   15310: gap of unknown length
*   15311   18683: contig of 3373 bp in length
*   18684   18783: gap of unknown length
*   18784   23200: contig of 4417 bp in length
*   23201   23300: gap of unknown length
*   23301   27524: contig of 4224 bp in length
*   27525   27624: gap of unknown length
*   27625   32318: contig of 4694 bp in length
*   32319   32418: gap of unknown length
*   32419   37205: contig of 4787 bp in length
*   37206   37305: gap of unknown length
*   37306   42143: contig of 4838 bp in length
*   42144   42243: gap of unknown length
*   42243   47622: contig of 5379 bp in length
*   47623   47722: gap of unknown length
*   47723   54326: contig of 6604 bp in length
*   54327   54426: gap of unknown length
*   54427   60002: contig of 5576 bp in length
*   60003   60102: gap of unknown length
*   60102   66820: contig of 6718 bp in length
*   66821   66920: gap of unknown length
*   66921   74094: contig of 7174 bp in length
*   74095   74194: gap of unknown length
*   74195   81658: contig of 7464 bp in length
*   81659   81758: gap of unknown length
*   81759   87920: contig of 6162 bp in length
*   87921   88020: gap of unknown length

```
*     88021    94235: contig of 6215 bp in length
*     94236    94335: gap of unknown length
*     94336   102294: contig of 7959 bp in length
*    102295   102394: gap of unknown length
*    102395   111827: contig of 9433 bp in length
*    111828   111927: gap of unknown length
*    111928   119373: contig of 7446 bp in length
*    119374   119473: gap of unknown length
*    119474   130756: contig of 11283 bp in length
*    130757   130856: gap of unknown length
*    130857   143566: contig of 12710 bp in length
*    143567   143666: gap of unknown length
*    143667   156273: contig of 12607 bp in length
*    156274   156373: gap of unknown length
*    156374   174943: contig of 18570 bp in length
*    174944   175043: gap of unknown length
*    175044   203094: contig of 28051 bp in length
*    203095   203194: gap of unknown length
*    203195   204665: contig of 1471 bp in length
*    204666   204765: gap of unknown length
*    204766   205864: contig of 1099 bp in length
*    205865   205964: gap of unknown length
*    205965   207143: contig of 1179 bp in length
*    207144   207243: gap of unknown length
*    207244   208378: contig of 1135 bp in length.

FEATURES             Location/Qualifiers
     source          1..208378
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-795C1"
     misc_feature    1..1933
                     /note="assembly_name:Contig10"
     gap             1934..2033
                     /estimated_length=unknown
     misc_feature    2034..4225
                     /note="assembly_name:Contig11"
     gap             4226..4325
                     /estimated_length=unknown
     misc_feature    4326..6989
                     /note="assembly_name:Contig12"
     gap             6990..7089
                     /estimated_length=unknown
     misc_feature    7090..9679
                     /note="assembly_name:Contig13"
     gap             9680..9779
                     /estimated_length=unknown
     misc_feature    9780..12591
                     /note="assembly_name:Contig14"
     gap             12592..12691
                     /estimated_length=unknown
     misc_feature    12692..15210
                     /note="assembly_name:Contig15"
     gap             15211..15310
                     /estimated_length=unknown
     misc_feature    15311..18683
                     /note="assembly_name:Contig16"
     gap             18684..18783
                     /estimated_length=unknown
     misc_feature    18784..23200
                     /note="assembly_name:Contig17"
     gap             23201..23300
                     /estimated_length=unknown
     misc_feature    23301..27524
                     /note="assembly_name:Contig18"
     gap             27525..27624
                     /estimated_length=unknown
     misc_feature    27625..32318
                     /note="assembly_name:Contig19"
     gap             32319..32418
                     /estimated_length=unknown
     misc_feature    32419..37205
                     /note="assembly_name:Contig20"
     gap             37206..37305
                     /estimated_length=unknown
     misc_feature    37306..42143
                     /note="assembly_name:Contig21"
     gap             42144..42243
                     /estimated_length=unknown
     misc_feature    42244..47622
                     /note="assembly_name:Contig22"
     gap             47623..47722
                     /estimated_length=unknown
     misc_feature    47723..54326
                     /note="assembly_name:Contig23"
     gap             54327..54426
                     /estimated_length=unknown
     misc_feature    54427..60002
                     /note="assembly_name:Contig24"
     gap             60003..60102
                     /estimated_length=unknown
     misc_feature    60103..66820
                     /note="assembly_name:Contig25"
     gap             66821..66920
                     /estimated_length=unknown
     misc_feature    66921..74094
                     /note="assembly_name:Contig26"
     gap             74095..74194
                     /estimated_length=unknown
     misc_feature    74195..81658
                     /note="assembly_name:Contig27"
     gap             81659..81758
                     /estimated_length=unknown
     misc_feature    81759..87920
                     /note="assembly_name:Contig28"
     gap             87921..88020
                     /estimated_length=unknown
     misc_feature    88021..94235
                     /note="assembly_name:Contig29"
```

```
Query Match         94.6%;   Score 2604.1;   DB 3;   Length 208378;
Best Local Similarity 99.0%;  Pred. No. 0;
Matches 2700;  Conservative  0;  Mismatches  19;  Indels  9;  Gaps  8;

Qy      1 GCCAGAAGCAGCTCAGCTTGGCAAGGTGGGAGATGACTGCTGTTCCCTTCGCATTGG     60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db 122926 GCCAGAAGCAGCTCAGCTTGGCAAGGTGGGAGATGACTGCTGTTCCCTTCGCATTTGG 122867

Qy     61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122866 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 122807

Qy    121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180
           ||||||||||| |||||||||||||||||||||||||||||| |  |  |||||||  |
Db 122806 GAGTGGATGAAGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAGTGGGACTGGGCAAGGC 122747

Qy    181 CCCCAGAGAGGGCAAGGTCTCTAGAGGGGGTCTCCCACATGACTGGCTTCACACAGGCA  240
           ||||||| ||||||||||||||||||| |||||||||||||||||||||||||||||||
Db 122746 CCCCAGAGCAGGGCAAGGTCTCTAGAGGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 122687

Qy    241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTTCCAGGGTTGCAGGTAGGAAAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122686 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTTCCAGGGTTGCAGGTAGGAAAT 122627

Qy    301 GTTGTACCCTCTTCTGATTGCCACCTCCTTCCATGCCCTTAGGGACACAGGGCTTGAG   360
           ||| ||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db 122626 GTTTGACCCTCTTCTGATTGCCACCTCCTTCCATGCCCCTTAGGGACACAGGGCTTGAG 122567

Qy    361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACGCGCCCAGGGGCACCCT  420
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||  |||
Db 122566 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCCGCCCAGGGG-ACCCT 122508

Qy    421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGTCCTCCTGGCTT  480
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db 122507 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 122448
```

(handwritten: page 3 of 6)

```
Qy    1561 CGTGGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCCAGCTCACACTCT 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121373 CGTGGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCCAGCTCACACTCT 121314

Qy    1621 TGTCCTGGAGGGGCAGCCTGATCTCACCTCACCTTGGGGACTGGGACTGGGGACCTT 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121313 TGTCCTGGAGGGGCAGCCTGATCTCACCTCACCTTGGGGACTTGGGGACTGAGGACCTT 121254

Qy    1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121253 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 121194

Qy    1741 AGGGACTGCTCACAGTGCCTCAGCTGGCCTCAGCTTGGTGGTTTGACGGCCGCCCCCGGCCCTCC 1800
           ||||||||||||||||||||||||||||||||||||||||||          |||||||||||
Db  121193 AGGGACTGCTCACAGTGCCTCAGCTGGCCTCAGCTTGGTGGTTTGACGGGCCG--CCCCGGCCCTCC 121135

Qy    1801 ATAAGGGTATCCTGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121134 ATAAGGGTATCCTGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 121075

Qy    1861 TGGAAGGAGTCCCACGGGAGTGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCCTGTG 1920
           ||||||||||||||||||||||| ||||||||          |||||||||||||||||||
Db  121074 TGGAAGGAGTCCCACGGGAGATGGGGTCCGAGGTCCGGCTGT-GCCATCCAGCCCCGCTGTG 121016

Qy    1921 GCTTGTCCAGCCTCGTGCGACCCCTGGTCTTCACTCAGGGCAGCAGCAGCAGCCACTG 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121015 GCTTGTCCAGCCTCGTGCGACCCCTGGTCTTCACTCAGGGCAGACAGCAGCAGCCACTG 120956

Qy    1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120955 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 120896

Qy    2041 GTGTTCTGCGATTGGTCAGCTTCCACTCGATCCTCCTAAAGCAATGGGGAGGCCCC 2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120895 GTGTTCTGCGATTGGTCAGCTTCCACTCGATCCTCCTAAAGCAATGGGGAGGCCCC 120836

Qy    2101 CACTAGCCCAGTTTCAGGAGTCAACTGGGACAACTGGGAGTTAGAGCAGTTAGATGGGGCCCAGGGTCCCACAGC 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120835 CACTAGCCCAGTTTCAGGAGTCAACTGGGACAACTGGGAGTTAGATGGGGCCCAGGGTCCCACAGC 120776

Qy    2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCGGTGTCCACTGCGATCCCACTTGCAGAT 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120775 TACTGATGGCCCGAGCCAGGTTGAGCTTCCGGTGTCCGATCCCACTTGCAGAT 120716

Qy    2221 CTCATGCTCTCAGATAGGTGGGACAGTCTTTTGTCACAGTTCTTGGCGGCTTCGTCCTGGGTCCTGAGG 2280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120715 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTGTCACAGTTCTTGGCGGCTCGGCCTTGTGTTCCTGAGG 120656

Qy    2281 CCTCATTGCTGCGCTGGGTGTCCTCTGCTGGGAAAGCTTTGCGGGCTTTGTGTTAAC 2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120655 CCTCATTGCTGCGCTGGGTGTCCTCTGCTGGGAAAGCTTTGCGGGCTTTGTGTTAAC 120596

Qy    2341 CACAGAGAGAAGGGGACTGTTTGGGGTGCCTCTCCGCAGCCTCCCGTGTGCTGGGTGGAA 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120595 CACAGAGAGAGAGGGGACTGTTTGGGGTGCCTCCGCAGCCTCCCGTGTGCTGGGTGGAA 120536

Qy    2401 GCACGGTTACTGTGTTCTCTAATGTTCATGGTTCATGTTCTAAAGATGTA 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120535 GCACGGTTACTGTGTTCTCTAATGTTCATGATTTCATTCTTTCTAAAGATGTA 120476

Qy    2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120475 ACCTCCACACCTTTCTCCAGATTGGGTGACTTTTTCTAAAGGTGTGGGAGTATCTGTC 120416

Qy    2521 GGGGTGGTGCCATGGTTCAGGTGGGTCTCCTGGGAGGTGCTGGGAGGTGGGCGT 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120415 GGGGTGGTGCCATGGTTCAGGTGGGTCTCCTGGGAGGTGCTGGGAGGTGGGCGT 120356

Qy    2581 TGAGCTCAAAGTTGTCCTACTGCACTGTTTTTGTACCTGAAATAAAGCATATTTGCACT 2640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  120355 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTGCACT 120296

Qy    2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
```

```
Qy     481 TCCTCCCTCCCCAATCTATGGTCACAGCTAACAGATCTGAGGGCAACTGTGTGCTAG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122447 TCCTCCCTCCCCAATCTATGGTCACAGCTAACAGATCTGAGGGCAACTGTGTGCTAG 122388

Qy     541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122387 TGGCCAGGGCTGCACCTGCCATCCCCGGCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 122328

Qy     601 TCCTTAGGAAGTAGCTCTGAGGCATGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122327 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 122268

Qy     661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCCGATGGGGAAA 720
           |||||||||||||||||||||||||||||          ||||||||||||||||||||||
Db  122267 TAAGGTGGGGAAGGACGGTCAGTGCTTGG--CCCCAGCTGCCAGCCTGGCCGATGGGGAAA 122210

Qy     721 CCAAAACATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCCCATGCCCTTCGTCCT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122209 CCAAAACATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCCCATGCCCTTCGTCCT 122150

Qy     781 GAGGAGCCCTGAGGTGGCAGCAGGGGCCAGGGAGTTTTCAGGCCTTCATCAAAGAGA 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  122149 GAGGAGCCCTGAGGTGGCAGCAGGGGCCAGGGAGTTTTCAGGCCTTCATCAAAGAGA 122090

Qy     841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900
           ||||||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Db  122089 ACAACATCCTCA-TGCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 122031

Qy     901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960
           ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  122030 CTTGTCAGCTAGCAT-CATCGTGCGCACCCCTGGCCCACTGGCTGTTTATGCCACTGATTTA 121972

Qy     961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTCTCCCCATCACAAAAAAAAA 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121971 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTCTCCCCATCACAAAAAAAAA 121912

Qy    1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAATATGTACTCAATTATTGTATTTTG 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121911 AATTCTTATTTTTAGTAGACATGTATTTTACCGTCTAGCCCGTTATTTGTATTTTG 121852

Qy    1081 GATTTTATCAATTAAAAATTGTGGAAATTTGTTGCTCTTACGGCCAACATAATATTGAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121851 GATTTTATCAATTAAAAATGTGGAAATTGTTTGCTCTTACGGCCAACATAATATTGAT 121792

Qy    1141 TTTGCCCTTGGCTCTGAAAGCCAAAATATTACCGGTCTAGCCCGTTACAGAAAAAGTC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121791 TTTGCCCTCTT-GTCTGAAAGCCAAAATATTACCGGTCTAGCCCGTTACAGAAAAAGTC 121733

Qy    1201 TGCTGACTACTGAGCCAGCAGACCTCCATTCACTCCCTGTTGGATTATTTAAAGAAAGC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121732 TGCTGACTACTGAGCCAGCAGACCTCCATTCACTCCCTGTTGGATTATTTAAAGAAAGC 121673

Qy    1261 CTCAGACAGTAAGCCCCCCCCTTTTTTAAAAGAATAAAAATGACTTGGTTGCGCTTGGAAGCAGG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121672 CTCAGACAGTAAGGGCCCCTTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 121613

Qy    1321 GGAAGCATTCAGATGAGCGGTTTCGCATTAACCCTGCCATCACGCATCTCGTGTCCTG 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121612 GGAAGCATTCAGATGAGCGGTTTCGCATTAACCCTGCCATCACGCATCTCGTGTCCTG 121553

Qy    1381 TGTGGCTGGCGAGCCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGCTTCCTGCAGAGTCCA 1440
           ||||||||||||||||          ||||||||||||||||||||||||||||||||||
Db  121552 TGTGGCTGGCGAG-CCCCCCTTGGAAGTTCTGGTGCTTCAGCTGCTTCCTGCAGAGTCCA 121494

Qy    1441 CCCCGCCTCGTGTGGGAATGCAGAGCCCTTTGCTTCTTCGCGGCCCTGCTTCCTGT 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121493 CCCCGCCTCGTGTGGGAATGCAGAGCCCTTTGCTTCTTCGCGGCCCTGCTTCCTGT 121434

Qy    1501 TCCTGGGACCCCGCTGGCCTTTGGTCATCCCCTGGGCAGGTCCCTCAGGGTTGATG 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121433 TCCTGGGGACCCCGCTGGCCTTTGGTCATCCCCTGGCCAGGGTCCCTCAGGGTTGATG 121374
```

Page 4 of 4

```
Db      120295  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 120236
Qy        2701  TGCAAATAAAACTCACATTTGTAAGAAA 2728
                ||||||||||||||||||||||||||||
Db      120235  TGCAAATAAAACTCACATTTGTAAGAAA 120208
```

RESULT 2
AC079988/c
LOCUS   AC079988   187722 bp   DNA   linear   HTG 11-OCT-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-79SC1, *** SEQUENCING IN
            PROGRESS ***, 57 unordered pieces.
ACCESSION   AC079988
VERSION   AC079988.2 GI:10765092
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 187722)
AUTHORS   Waterston,R.H.
TITLE   The sequence of Homo sapiens clone
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 187722)
AUTHORS   Waterston,R.H.
TITLE   Direct Submission
JOURNAL   Submitted (20-SEP-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT   [WARNING] On Oct 13, 2000 this sequence was replaced by a newer
            version gi:10800346.
            On Oct 11, 2000 this sequence version replaced gi:10198585.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------------- Project Information ----------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 57 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *   1   1302: contig of 1302 bp in length
            *   1303   1402: gap of unknown length
            *   1403   2789: contig of 1387 bp in length
            *   2790   2889: gap of unknown length
            *   2890   4094: contig of 1205 bp in length
            *   4095   4194: gap of unknown length
            *   4195   5356: contig of 1162 bp in length
            *   5357   5456: gap of unknown length
            *   5457   6668: contig of 1212 bp in length
            *   6669   6768: gap of unknown length
            *   6769   7867: contig of 1099 bp in length
            *   7868   7967: gap of unknown length
            *   7968   9464: contig of 1497 bp in length
            *   9465   9564: gap of unknown length
            *   9565   11076: contig of 1512 bp in length
            *   11077   11176: gap of unknown length
            *   11177   12520: contig of 1344 bp in length
            *   12521   12620: gap of unknown length
            *   12621   14052: contig of 1432 bp in length
            *   14053   14152: gap of unknown length
            *   14153   15875: contig of 1723 bp in length
            *   15876   15975: gap of unknown length
            *   15976   17390: contig of 1415 bp in length
            *   17391   17490: gap of unknown length
            *   17491   19569: contig of 2079 bp in length
            *   19570   19669: gap of unknown length
            *   19670   20689: contig of 1020 bp in length
            *   20690   20789: gap of unknown length
            *   20790   22686: contig of 1897 bp in length
            *   22687   22786: gap of unknown length
            *   22787   25402: contig of 2616 bp in length
            *   25403   25502: gap of unknown length
            *   25503   27972: contig of 2370 bp in length
            *   27873   27972: gap of unknown length
            *   27973   29981: contig of 2009 bp in length
            *   29982   30081: gap of unknown length
            *   30082   32024: contig of 1943 bp in length
            *   32025   32124: gap of unknown length
            *   34083   34183: contig of 1959 bp in length
            *   34084   34183: gap of unknown length
            *   34184   37046: contig of 2863 bp in length
            *   37047   37146: gap of unknown length
            *   37147   39347: contig of 2201 bp in length
            *   39348   39447: gap of unknown length
            *   39448   42510: contig of 3063 bp in length
            *   42511   42610: gap of unknown length
            *   42611   45752: contig of 3142 bp in length
            *   45753   45852: gap of unknown length
            *   45853   48296: contig of 2444 bp in length
            *   48297   48396: gap of unknown length
            *   48397   50281: contig of 1885 bp in length
            *   50282   50381: gap of unknown length
            *   50382   52838: contig of 2457 bp in length
            *   52839   52938: gap of unknown length
            *   52939   55722: contig of 2784 bp in length
            *   55723   55822: gap of unknown length
            *   55823   58127: contig of 2305 bp in length
            *   58128   58227: gap of unknown length
            *   58228   60861: contig of 2634 bp in length
            *   60862   60961: gap of unknown length
            *   60962   64066: contig of 3105 bp in length
            *   64067   64166: gap of unknown length
            *   64167   67176: contig of 3010 bp in length
            *   67177   67276: gap of unknown length
            *   67277   69558: contig of 2282 bp in length
            *   69559   69658: gap of unknown length
            *   69659   72835: contig of 3177 bp in length
            *   72836   72935: gap of unknown length
            *   72936   75536: contig of 2601 bp in length
            *   75537   75636: gap of unknown length
            *   75637   79373: contig of 3737 bp in length
            *   79374   79473: gap of unknown length
            *   79474   82535: contig of 3062 bp in length
            *   82536   82635: gap of unknown length
            *   82636   85680: contig of 3045 bp in length
            *   85681   85780: gap of unknown length
            *   85781   89885: contig of 4105 bp in length
            *   89886   89985: gap of unknown length
            *   89986   94532: contig of 4547 bp in length
            *   94533   94632: gap of unknown length
            *   94633   98426: contig of 3794 bp in length
            *   98427   98526: gap of unknown length
            *   98527 102557: contig of 4031 bp in length
            *   102558 102657: gap of unknown length
            *   102658 108036: contig of 5379 bp in length
            *   108037 108136: gap of unknown length
            *   108137 113024: contig of 4888 bp in length
            *   113025 113124: gap of unknown length
            *   113125 116427: contig of 3303 bp in length
            *   116428 116527: gap of unknown length
            *   116528 120156: contig of 3529 bp in length
            *   120157 120156: gap of unknown length
            *   120057 125167: contig of 5011 bp in length
            *   125168 125267: gap of unknown length
            *   125268 131332: contig of 6065 bp in length
            *   131333 131432: gap of unknown length
            *   131433 136663: contig of 5231 bp in length
            *   136664 136763: gap of unknown length
            *   136764 141035: contig of 4272 bp in length
```